Essays on Experimental Economics for the Environment and Economics of Privacy

Doctoral Thesis

to be awarded the degree

Doctor of Economics (Dr. rer. pol.)

submitted by

Janis Cloos, M.Sc.

from Giessen

approved by the Faculty of Energy and Economic Sciences,
Clausthal University of Technology,

Date of oral examination

July 21, 2021

Dean

Prof. Dr. mont. Leonhard Ganzer

Chairperson of the Board of Examiners

Prof. Dr. mont. Leonhard Ganzer

Supervising Tutor

Prof. Dr. rer. pol. Matthias Greiff

Reviewer

Prof. Dr. sc. pol. Roland Menges

Acknowledgments

This dissertation was written during my time at the Clausthal University of Technology and has benefited in various ways from the support of a large number of people.

I would like to express my greatest thanks to my supervisor Matthias Greiff. Over the past years, Matthias has supported my scientific and personal development with great commitment. From the very beginning, he motivated me to develop my own research ideas and to continuously expand and improve my skills. Already during my master's studies, he played a key role in deepening my interest for science. All papers included in this dissertation have benefited greatly from Matthias' co-authorship or from his valuable comments.

I would also like to sincerely thank my second supervisor Roland Menges. Several papers in this dissertation have benefited from Roland's enriching comments or his co-authorship. In addition, the exciting soccer matches with Roland have helped me to maintain my physical fitness in spite of spending most of the work on this dissertation at the desk.

My deepest appreciation goes to Mathias Erlei, who sadly passed away much too early and at whose chair I had the privilege to work for a few months in 2018.

I am grateful to all my co-authors, especially Hannes Rusch, Björn Frank, and Jacob Wehrle. I have greatly benefited from the professional advice provided by Heike Schenk-Mathes and Christoph Bühren. I would also like to thank my colleagues Julien Geissmar, Maximilian Vincent Reimer, Philipp Balkwitz and Janis Kesten-Kühne. Many colleagues not mentioned by name also deserve my sincere gratitude. In addition, I would like to offer my special thanks to Astrid Ludwig. Her competent support in numerous tasks has contributed considerably to the completion of this dissertation. The suggestions of Merindah Loessl have improved the text of this dissertation substantially.

I am deeply grateful to my partner and co-author Svenja for her loving support and patience. Finally, I am indebted to my parents Bärbel and Harry, to whom I dedicate this dissertation, as well as my family and friends, for their continuous support and for always motivating me to get excited about the fascinating aspects of life.

Abstract

In the 21st century, two main challenges for economic research are to propose effective solutions to shape the digital transformation and mitigate human-induced climate change. Research on digital transformation is closely linked to various privacy-related issues, which mostly relate to the preferences and decisions of individuals. In contrast, climate change research examines which factors impede effective cooperation among multiple individuals and investigates how common goals, such as limiting climate change, can be achieved.

The link between economics of privacy and environmental economics is that many digital technologies have the potential to generate positive externalities that can contribute to the provision or maintenance of public goods. However, in many cases these digital technologies are characterized by the fact that their use requires the disclosure of personal information. The potential success of these technologies and institutional mechanisms therefore largely depends on social acceptance towards these technologies and institutional mechanisms.

Each paper in this cumulative dissertation contributes to the broader question of how economic experiments can contribute to evaluate and potentially increase the efficiency of institutions and technologies that can provide or maintain public goods. The first paper investigates whether the publication process of journals in the field of experimental economics can potentially be improved. The remaining five papers focus directly or indirectly on different but related public goods problems which are closely linked to privacy or environmental issues. Methodologically, the six papers share the feature that they either directly apply the experimental method for their individual research questions or use the results of experimental literature to derive hypotheses and explain empirical outcomes in specific privacy-related contexts.

In the field of privacy, the dissertation identifies factors that influence data sharing in several smartphone apps from key industries of the digital transformation and on employer review platforms. In the area of environmental economics, the first paper proposes an institutional mechanism that can increase the willingness to contribute to recycling systems, and the second paper shows that the ability to exploit a public good can impede cooperation to mitigate climate change.

Zusammenfassung

Im 21. Jahrhundert bestehen zwei Hauptherausforderungen der ökonomischen Forschung darin, effektive Lösung für die Gestaltung der digitalen Transformation und für die Eindämmung des menschengemachten Klimawandels aufzuzeigen. Die Forschung zur digitalen Transformationen ist eng mit verschiedenen Datenschutz- (oder Privatsphäre-)relevanten Fragestellungen verbunden, die sich vorwiegend auf die Präferenzen und Entscheidungen von Einzelpersonen beziehen. Im Gegensatz dazu befasst sich die Forschung zum Klimawandel damit, welche Faktoren eine effektive Kooperation zwischen mehreren Individuen erschweren und wie gemeinsame Ziele, wie die Begrenzung des Klimawandels, erreicht werden können.

Die Verbindung zwischen Datenschutz- und Umweltökonomie besteht darin, dass viele digitale Technologien das Potential haben, positive externe Effekte zu erzeugen, die zur Bereitstellung oder Erhaltung öffentlicher Güter beitragen können. Oftmals sind diese digitalen Technologien jedoch dadurch gekennzeichnet, dass ihre Nutzung die Offenlegung persönlicher Informationen erfordert. Der potentielle Erfolg dieser Technologien und institutionellen Mechanismen hängt daher weitgehend von der gesellschaftlichen Akzeptanz gegenüber diesen Technologien und institutionellen Mechanismen ab.

Jeder Artikel in dieser kumulativen Dissertation leistet einen Beitrag zu der übergeordneten Fragestellung, inwiefern ökonomische Experimente dazu beitragen können, die Effizienz von Institutionen und Technologien, die öffentliche Güter bereitstellen oder erhalten können, zu evaluieren und potentiell zu steigern. Im ersten Artikel wird untersucht, ob der Publikationsprozess von Fachzeitschriften im Bereich der experimentellen Ökonomik verbessert werden kann. Die weiteren fünf Artikel befassen sich direkt oder indirekt mit unterschiedlichen, aber miteinander verbundenen Problemstellungen zu öffentlichen Gütern, die eng mit Fragen zum Datenschutz oder Umweltfragen verbunden sind. Methodisch sind die sechs Artikel dadurch gekennzeichnet, dass sie die experimentelle Methode entweder direkt für ihre individuellen Forschungsfragen anwenden oder die Ergebnisse der experimentellen Literatur nutzen, um Hypothesen abzuleiten und empirische Ergebnisse in spezifischen Datenschutz-relevanten Kontexten zu erklären.

Im Bereich des Datenschutzes werden in der Dissertation Faktoren identifiziert, die die Weitergabe von Daten in verschiedenen Smartphone-Apps aus Schlüsselindustrien der digitalen Transformation und auf Arbeitgeberbewertungsplattformen beeinflussen. Im Bereich der Umweltökonomie wird im ersten Artikel ein institutioneller Mechanismus vorgeschlagen,

der die Bereitschaft erhöhen kann, zu Recyclingsystemen beizutragen und im zweiten Artikel wird gezeigt, dass die Möglichkeit, ein öffentliches Gut auszubeuten, die Kooperation zur Eindämmung des Klimawandels erschweren kann.

Contents

List of Figures

List of Tables

1. Introduction

The 21st century can be described as the age of information (Acquisti et al., 2016) or the age of big data (Einav and Levin, 2014). The continuous expansion of fast internet connections and the widespread use of digital devices have led to digital applications being used in almost every area of life, especially in developed societies. Unlike in the pre-internet era, nowadays, numerous activities and economic decisions of individuals, companies, and other organizations are collected, stored, and shared with different actors through digital applications. As Acquisti et al. (2016) illustrate, these data sharing and the associated continuously evolving opportunities for data use and analysis hold many benefits for individuals, firms, and societies. Smart meters (intelligent digital electricity meters), for example, enable households to optimize their individual electricity consumption behavior (Zheng et al., 2013). In addition to the financial benefits to these households, an optimized electricity consumption behavior may also result in positive externalities, for example, in the form of reduced CO_2 emissions and lower electricity prices in the long run. However, the almost permanent monitoring and collection of data does not come without risks. As an example, Acquisti et al. (2016) emphasize the particular problem of data being stored permanently. Even years later, individuals may still suffer negative consequences from information that was generated and digitally stored in the past.

The benefits and risks of privacy-related digital technologies therefore generate a wide range of trade-offs that have been the subject of economic research. Economic research methods, such as experiments, are well suited to contribute to a better understanding of privacy-related decision making by individuals and organizations. The insights gained through economic research on privacy can help governments, companies, and consumer advocates to adequately address privacy-related issues in the development of laws, regulations, products, and services.

The development of new digital technologies and business models raises many privacy-related research questions that constantly pose new challenges for both economic research and research in other disciplines. In contrast to research on privacy issues, research on climate change is of even greater importance for economics and science in general, as it not only massively affects people living today but also future generations, other species, and ecosystems. William D. Nordhaus (2019), Nobel Laureate of 2018[1], describes climate change as the ultimate challenge for economics. Although climate change is the result of economic activities, there is currently

[1] "For integrating climate change into long-run macroeconomic analysis" (The Nobel Foundation, 2018).

a comparatively small amount of economic research on climate change. As Oswald and Stern (2019) show, the number of papers on climate change published in top economics journals has been extremely low[2].

Human-induced climate change, which is caused by greenhouse gas emissions (predominantly CO_2), is increasingly becoming visible in many parts of the world, e.g., through severe droughts, rising sea levels, a decline in biodiversity, or a loss of ecosystems (IPCC, 2018). The climate can be described as a global public good, which is characterized by the fact that it can be consumed simultaneously by different individuals (non-rivalry) and that no one can be excluded from consumption (non-excludability). Global warming is a negative externality that results from greenhouse gas emissions and which reduces the value of climate as a public good. The global nature of this externality makes it difficult to combat climate change at the level of individual states or via pure market solutions. Consequently, a key task of environmental economists is to propose pathways that enable global intergovernmental coordination to combat climate change (Nordhaus, 2019). At the local (country) level, environmental economists can, e.g., evaluate the efficiency of climate-friendly technologies and institutional mechanisms and identify incentives to increase societal acceptance, and adaptation of these technologies and mechanisms.

The relationship between economics of privacy and environmental economics is that many climate-friendly technologies and actions require digital sharing of the data they generate. An appropriate example is the smart meter mentioned above. A smart meter can enable a household with a photovoltaic system to switch on electricity-intensive household devices precisely when the power supplied via the photovoltaic system is particularly high. However, besides these advantages, smart meters can also give rise to privacy concerns among consumers because, unlike previous analog electricity meters, they regularly transmit consumption data digitally to energy providers via a so-called gateway. Another example can be found in the area of recycling. Privacy concerns may be one reason why consumers prefer to keep broken or unused digital devices such as smartphones, laptops, or tablets because of the data stored on them instead of recycling them appropriately. To ensure that large parts of a society adapt

[2] As of August 2019, there have been 0 articles published on climate change in *The Quarterly Journal of Economics*, although about 4,700 articles have been published in the journal since its foundation in 1886. The number of related articles in other top economics journals is also extremely low (19 in the *American Economic Review*, 8 in the *Journal of the European Economic Association*, 3 in *The Review of Economic Studies*, 2 in *Econometrica*, and 9 in the *Journal of Political Economy*).

climate-friendly technologies, such as smart meters, it is therefore necessary to identify potential privacy concerns regarding these technologies and to address them accordingly.

This dissertation examines issues on the economics of privacy and environmental economics. The six papers in this dissertation are connected by the feature that they directly apply the experimental method for their individual research questions or that they use the results of experimental literature to derive hypotheses and explain empirical outcomes in specific contexts. Each paper thus contributes to the broader question of how economic experiments can contribute to evaluate and potentially increase the efficiency of institutions[3] and technologies that can provide or maintain public goods.

The remainder of this dissertation is structured as follows: In chapter 1.1, I briefly discuss the experimental method in order to provide some background information for the experimental methods used in the individual chapters. In chapter 1.2, I summarize the individual chapters and discuss their interrelation. My individual contribution to each chapter is explained in chapter 1.3. Chapters 2 to 7 contain the individual papers and chapter 8 gives a conclusion and outlook.

1.1 The experimental method in economics

In this section, I briefly discuss the experimental method and explain the types of experiments used in the papers of this dissertation.

According to Smith (1991), economic experiments rely on the three components environment, institution and behavior. The environment specifies the initial endowments, costs, and preferences that are intended to motivate participants' decisions in the experiment. The environment is controlled by monetary incentives which are supposed to reflect the ratio between costs and preferences in a desired way. The institution describes the rules and terminology of information exchange within an experiment as well as the possible moves and the potential outcomes of specific decisions. In laboratory experiments and some types of field experiments, these institutional characteristics are explained in the experimental instructions. In experiments on auction markets (see Kagel and Levin, 2010), e.g., the instructions clearly specify how bids and offers will be presented and how binding contracts will be formed. The

[3] There is no uniform definition for the term institution. This dissertation follows the definition of Erlei et al. (2016), according to which institution describes a contract or contract system, or a rule or rule system, as well as the related enforcement mechanisms.

behavior of the participants in the experiment is the result of the interaction of the environment and the institution. While taking relevant control variables into account, experimental economists can use experiments to draw conclusions about which changes in the environment or the institution cause changes in participants' observed behavior.

The use of experiments in economics constitutes a comparatively young field of research compared to other disciplines. The establishment of experimental methods dates back to the middle of the 20th century (Smith, 2010; Davis and Holt, 1993; Roth, 1993). Today, experimental economics is an established field of research within the discipline. This is underlined by the widespread use of experiments to address economic problems (Nikiforakis and Slonim, 2019) and by the Nobel Prizes awarded to Vernon Smith and Daniel Kahneman in 2002[4], to Richard Thaler in 2017[5], and to Abhijit Banerjee, Esther Duflo, and Michael Kremer in 2019[6].

In the economic literature, there are different concepts of what constitutes an experiment. Shadish et al. (2002) describe an experiment as a study in which certain interventions have been intentionally introduced by the experimenter in order to examine the effects of those interventions. In contrast, Czibor et al. (2019) use a broader definition and describe an experiment as a study in which primary data are collected in a controlled environment.

There are also divergent views on what characterizes laboratory and field experiments. According to the broad definition of Friedman et al. (1994), laboratory experiments are used to collect data in a controlled environment that is adapted to the context of interest. In contrast, data generation in field experiments takes place in a natural environment. Harrison and List (2004) use a more detailed distinction of different types of (field) experiments. According to Harrison and List (2004), the characteristic of a laboratory experiment is that the behavior of participants, in this case university students, is studied in an artificial environment. For example, Greiff and Paetzel (2020) let university students play a repeated public goods game in which each player has to evaluate the contributions of her partner in order to gain insights into the mechanisms of reputation systems. Following Harrison and List (2004), an artefactual field experiment differs from a laboratory experiment only in that the participants are drawn

[4] "For having integrated insights from psychological science, especially concerning human judgement and decision-making under uncertainty" (Kahneman). "For having established laboratory experiments as a tool in empirical economic analysis, especially in the study of alternative market mechanisms" (Smith) (The Nobel Foundation, 2002).

[5] "For his contributions to behavioral economics" (The Nobel Foundation, 2017).

[6] "For their experimental approach to alleviating global poverty" (The Nobel Foundation, 2019).

from the population of interest. Artefactual field experiments can therefore also be described as lab-in-the-field experiments. For example, Menges et al. (2005) recruited consumers in a shopping center to participate in an experiment investigating the impacts of warm glow, pure altruism, and crowding-out effects on consumers' demand for green electricity. As described by Harrsion and List (2004), framed field experiments and natural field experiments use the population of interest and natural environments for their research questions. However, I will not discuss these types of experiments, as they are not part of this dissertation.

While laboratory and field experiments are used for purely scientific reasons, there are data that are produced through economic activities (Friedman et al., 1994). Although the primary intention behind these economic activities is not to generate research data, in some cases, such data can be described as a natural experiment. According to Shadish et al. (2002), a natural experiment is characterized by some natural event that results in a particular group of people being exposed to a special condition, whereas another (control) group is not exposed to that condition. Unlike in laboratory or field experiments, this special condition, which can potentially cause outcomes to differ between groups, cannot be manipulated by the researcher.

In the following, I briefly discuss the strengths and limitations of laboratory and artefactual field experiments, since these types of experiments are part of this dissertation. Laboratory and artefactual field experiments allow to look at human actions in controlled environments and, in this way, to determine in isolation what factors may facilitate or impede the occurrence of specific outcomes. Additionally, both types of experiments can be replicated comparatively easily (Davis and Holt, 1993). Camerer et al. (2016) examine the replicability of 18 economic experiments published in the *American Economic Review* and *The Quarterly Journal of Economics* between 2011 and 2014. For various replication indicators, the authors find replication rates between 61 and 78 %.[7] The authors conclude that economic experiments published in top economics journals have comparatively high replication rates.[8] However, a frequently raised objection to laboratory experiments with student participants is that these participants are WEIRD (western, educated, industrialized, rich, and democratic) (Henrich et al., 2010a, b). Since the representativeness of WEIRD samples is limited, it is highly questionable whether and to what extent results from WEIRD samples allow conclusions

[7] Camerer et al. (2019) find similarly high replication rates for 21 social science experiments published in the interdisciplinary journals *Nature* and *Science* between 2010 and 2015.

[8] A related study by the Open Science Collaboration (2015) shows significantly lower replication rates for psychological experiments. Out of 100 replicated experiments, only 36% showed significant effects, whereas significant effects were found in 97% of the original studies.

to be drawn on other populations. Because (artefactual) field experiments use the population of interest as participants, their central strength is that the question of transferability of results to other populations is often less problematic than in laboratory experiments (for detailed discussions see Czibor et al., 2019; Harrison and List, 2004).

Laboratory experiments and artefactual field experiments allow to control for a large number of variables. By accounting for control variables, different theories and their corresponding models can be evaluated specifically by the behavior observed in the laboratory. This is not possible in a natural environment where less accurate data collection capabilities prevent the evaluation of differing theories. Furthermore, situations that allow to test a particular theory are often difficult to find in the world outside the laboratory. In such cases laboratory experiments or artefactual field experiments are often the only applicable method for testing a theory (Davis and Holt, 1993).

Both types of experiments allow to make exogenous changes without altering the remaining characteristics of the experiment. In contrast, the most interesting variables in field experiments in natural environments are often endogenously predetermined. As a result, in natural environments, it is difficult to infer causal relationships, and at best, a correlation can be measured between the variables of interest (Falk and Fehr, 2003). Gadenne (2011) points out that for the appropriate choice between a natural and an artificial research environment, no universal statement can be made. Instead, the decisive factor for the choice of the appropriate research method should be in which environment the initial conditions of the theory under investigation are most likely to be found. Falk and Heckman (2009) specify this by emphasizing that the main problem is to isolate the causal effect of interest from other effects.

1.2 Description of the research topics and contextual connection of the individual papers

The dissertation is organized in three parts. Part I is entitled *Methodological standards and editorial favoritism in experimental economics* and contains one paper. The paper serves as a prelude to the main section as it empirically examines methodological aspects of laboratory experiments and therefore provides a good foundation for the five other papers. The paper empirically complements previous literature on the methodology of experimental economics (e.g., Greiff and Cloos, 2019; Fréchette and Schotter, 2015; Bardsley et al., 2010; Hertwig and Ortmann, 2001) and the literature on editorial favoritism in economics (Colussi, 2018; Brogaard et al., 2014; Medoff, 2003; Laband and Piette, 1994a).

Part II has the title *Empirical privacy and acceptance research on digital platforms and technologies* and includes three papers. Although all three papers follow their own research questions, each of them examines the effect of individual privacy preferences on economic decision making related to, e.g., the provision of public goods. The papers add to the extensive body of economic research on privacy that has emerged in the context of increasing digitalization since the early 2000s (for overviews, see Kokolakis, 2017; Acquisti et al., 2016, 2015; Smith et al., 2011).

Part III of the dissertation, which is entitled *Experimental environmental economics*, contains two papers that are part of the third-party funded project "Innovative acceptance research for sustainable development through gamification (GAME)". The papers address issues on recycling and climate change by examining how the provision of public goods is affected by the framing of the action sets of the actors involved. The paper on recycling extends experimental research on club goods and group identity (e.g., Chakravarty and Fonseca, 2017; Chen and Li, 2009) and recycling (e.g., Huhtala, 2009; Kaoursakis and Birol, 2008). The paper on climate change complements experimental research on the so-called collective risk social dilemma (Milinski et al. 2008; for an overview see Dannenberg and Tavoni, 2016). Table 1.1 shows the structure of the dissertation ordered by part, chapter, and the methods used. The two columns on the right list the number of my author points for the respective paper, as well as the status of the paper (as of April 16, 2021).

Part	Chapter	Method	Author points	Status
I. Methodological standards and editorial favoritism in experimental economics	2. Editorial favoritism in the field of laboratory experimental economics	Empirical analysis and theory	0.7	Under review at *The Economic Journal*; available as GSBE working paper (Maastricht University)
II. Empirical privacy and acceptance research on digital platforms and technologies	3. Is your privacy for sale? An experiment on the willingness to reveal sensitive information	Classroom experiment (laboratory experiment)	0.1	Published in *Games*
	4. Employer review platforms – do the rating environment and platform design affect the informativeness of reviews? Theory, evidence, and suggestions	Empirical analysis and theory (natural experiment)	1	Accepted at *mrev management revue*
	5. Acceptance of data sharing in smartphone apps from key industries of the digital transformation: A representative population survey for Germany	Online survey experiment (artificial field experiment)	0.5	Under review at *Technological Forecasting and Social Change*; available as MAGKS working paper
III. Experimental environmental economics	6. Recycling behavior of private households: an empirical investigation of individual preferences in a club good experiment	Classroom experiment (laboratory experiment)	0.16	Published in *Clean Technologies and Environmental Policy*
	7. Combating climate change: Is the option to exploit a public good a barrier for reaching critical thresholds? Experimental evidence	Laboratory experiment	0.75	Available as working paper

Table 1.1: Structure of the dissertation.

Following the typology of Harrison and List (2004), the experiments in this dissertation can be classified as laboratory experiments and artefactual field experiments. In addition to the laboratory experiment in chapter 7, I classify the two classroom experiments in chapters 3 and 6 as laboratory experiments since they were conducted with university students in artificial environments. The survey experiment[9] on the acceptance of data sharing in smartphone apps (chapter 5) is classified as an artefactual field experiment, since the participants were drawn from a representative sample of the German population and made their decisions in artificial environments. Chapter 4 of this dissertation contains research that can be broadly described as a natural experiment, as I explain below.

[9] The survey experiment deviates from Smith's (1991) definition with respect to the component environment, since the decisions within the survey are not payoff relevant for the participants. Consistent with related literature (e.g., John et al., 2011), we nevertheless refer to the survey as an experiment because participants are randomly assigned to two different treatments and the treatments differ systematically with respect to one component.

Part I of this dissertation includes the paper "Editorial favoritism in the field of laboratory experimental economics" (co-authored by Matthias Greiff and Hannes Rusch). In the paper, we study whether there is evidence for editorial favoritism in the field of experimental economics by examining all laboratory experiments published in the three journals *American Economic Review* (AER), *Experimental Economics* (EE), and *Journal of the European Economic Association* (JEEA) from 1998 to 2018. Editorial favoritism describes the phenomenon that submitted manuscripts from authors with a social tie to one or more editors of a journal (e.g., through a previous co-authorship) are favored for publication over submitted manuscripts by authors without these social ties. If editorial favoritism leads to lower-quality manuscripts from authors with social ties being accepted for publication rather than only the best-quality manuscripts, the public good of science will not be provided in an optimal way. Editorial favoritism can therefore be a serious problem both for trust in science and for scientific progress.

In the paper, we first use several ex ante proxies for the quality of laboratory experiments to analyze whether common methodological standards for conducting laboratory experiments exist across different regions (or experimental schools). Using the citations received by each paper in the years following publication as a measure for ex post quality, and controlling for ex ante quality proxies and further control variables, we examine whether there is evidence that authors with social ties to journal editors are favored over authors without such social ties. Our results show that there are no common methodological standards with regard to ex ante quality proxies. Several of these proxies are significantly higher for laboratory experiments conducted in Europe than for those conducted in the US. Laboratory experiments conducted in Europe also receive significantly more citations in the years following publication. We also find that essential details of experiments are often not reported in the papers which can complicate replication of the experiments. For the AER and the JEEA our results suggest that papers by authors with social ties to the editors are of lower quality than papers by authors without social ties. We thus find evidence that editorial favoritism in the field of laboratory experimental economics is an issue in these journals.

Part II of this dissertation contains three papers. The first paper „Is your privacy for sale? An experiment on the willingness to reveal sensitive information“ (co-authored by Björn Frank, Lukas Kampenhuber, Stephany Karam, Nhat Luong, Daniel Möller, Maria Monge-Larrain, Nguyen Tan Dat, Marco Nilgen, and Christoph Rössler) can be classified as experimental basic research. The paper includes an economic classroom experiment with a new experimental

approach to the so-called privacy paradox. The privacy paradox refers to people often revealing their personal data freely in real-life situations, despite stating in surveys that they care about the protection of their data. In the paper, we investigate whether there is a positive correlation between an index of privacy-related behavior on the internet and the willingness to accept the disclosure of personal data to other people. Furthermore, we investigate to what extent the willingness to disclose personal data is influenced by the concrete content of this data. According to our results, there is no conclusive relationship between the index of privacy-related behaviors and the willingness to accept. However, our results clearly show that the willingness to accept the disclosure of personal data to other people is influenced by the specific content of the personal data. Significantly higher amounts of money are demanded for more sensitive information than for less sensitive information. The findings of this paper are used in the two following papers to derive theoretical predictions in the investigation of practical problems related to privacy.

The second paper of part II has the title „Employer review platforms – do the rating environment and platform design affect the informativeness of reviews? Theory, evidence, and suggestions". In the paper, I use the employee reviews for 114 companies on the employer review platforms (ERPs) *Kununu* and *Glassdoor* to examine the extent to which the respective platform design and specific characteristics of the rating environment impact employee reviews. I distinguish the rating environment on ERPs from the well-studied rating environment on online marketplaces to highlight specific characteristics of ERPs. I explain that the ratings provided on ERPs can be seen as a public good and that ERPs can help to reduce information asymmetries between employees and employers and, thus, improve the job matching process in the long run. Based on the results of previous economic and psychological experiments, I identify different factors that could have a biasing influence on reviews. The results show that the reviews on *Glassdoor* are significantly better than on *Kununu* which can be explained by the larger amount of mandatory information on *Kununu*. Based on theoretical arguments and empirical analyses, I show that factors such as employee's awareness of their impact on a company's reputation and the perceived level of anonymity can also have a biasing influence on reviews. Drawing on these results, I suggest design changes that would allow the aggregate reviews to be presented in a more informative way and thus generate more value for the users and operators of such platforms.

The comparison of the reviews on the two platforms *Kununu* and *Glassdoor* can be classified as a natural experiment in the widest sense. The employees are exposed to different

environments on the platforms, but the researcher cannot influence these environments and the self-selection of the employees to the two platforms. However, it should be noted that it is quite possible that individual employees have rated the same company on both *Kununu* and *Glassdoor*. Based on the available data, however, no statement can be made about the extent to which the user groups on *Kununu* and *Glassdoor* overlap.

The third and final paper of part II is entitled „Acceptance of data sharing in smartphone apps from key industries of the digital transformation: A representative population survey for Germany" (co-authored by Svenja Mohr). The paper provides a link between economics of privacy (part II) and environmental economics (part III). In the paper, we use an online survey to investigate the acceptance of data sharing in various smartphone apps using hypothetical albeit realistic scenarios. The providers of the described apps belong to industries or areas that are strongly affected by the digital transformation. More specifically, we examine the extent to which acceptance towards data sharing in apps is influenced by the potential recipients, the collected information attributes, and the highlighted benefits of data sharing. We distinguish the highlighted benefits of data sharing in two treatments. While data sharing in treatment 1 is linked to monetary (or personal) benefits, data sharing in treatment 2 is linked to environmental (or public) benefits. Although we find no treatment effect in our results, we are able to identify significant differences in the acceptance values for different recipients and information attributes. A notable result of our survey is that there is a significant positive relationship between green consumption values and willingness to accept data sharing.

Part III of this dissertation contains two papers. In the first paper „Recycling behavior of private households: an empirical investigation of individual preferences in a club good experiment" (co-authored by Roland Menges, Matthias Greiff, Jacob Wehrle, Daniel Goldmann, and Lisa Rabe) we conduct an economic classroom experiment as a pilot study on recycling behavior. We model recycling behavior in the experiment as a public goods game in which players in groups of six decide individually over ten rounds how much of their initial endowment to contribute to a common pot (the public recycling system). The initial endowment reflects consumption which generates waste. The recycling of this waste is costly for the individual players but generates positive externalities from which all players in a group benefit. While in treatment 1, all players of a group can contribute to a common public recycling system, in treatment 2, players are divided into two clubs (high and low) and can contribute to a club recycling system in addition to the public recycling system. The division of the clubs was determined by a survey on the environmental attitudes of the players at the beginning of the

experiment. In the experiment, we use a within-subject design in which all players first participate in treatment 1 and then in treatment 2. Our results suggest that the introduction of clubs leads to higher total contributions in treatment 2 (public and club system) than in treatment 1 (public system). However, total contributions are significantly higher only in the club with stronger environmental attitudes.

The second paper of part III is titled "Combating climate change: Is the option to exploit a public good a barrier for reaching critical thresholds? Experimental evidence" (co-authored by Matthias Greiff). In the paper, we conduct a threshold public goods experiment with six treatments which is an extension of experiments on the so-called collective risk social dilemma (CRSD) by Milinski et al. (2011, 2008). In the experiment, each player is assigned to a group of six and has two accounts, called operating fund and endowment. Over the course of 20 rounds, players can contribute amounts from their operating fund to a joint group account. If the value of the group account exceeds a certain threshold after the 20 rounds, each player receives both his non-invested operating fund and her endowment. However, if a group falls short of the threshold, a treatment-dependent fixed percentage of the endowment is lost. In this case, each player only receives her non-invested operating fund and the non-lost part of the endowment. Unlike in previous variants of the CRSD, participants in our experiment can also withdraw (take) amounts from and contribute (give) amounts to the group account in three of the six treatments. In this way, we experimentally model the possibility of the exploitation of public goods which is a major issue in climate negotiations. In all treatments, we let the computer decide on the contributions in the first 10 rounds and, in this way, induce heterogeneity in the operating funds in four of the six treatments. Between treatments, we also vary the potential loss rate in order to allow for comparability of equilibria in give and give-take treatments.

In contrast to our hypotheses, results do not show that success rates in reaching the threshold are significantly lower in give-take treatments compared to pure give-treatments. However, in treatments with the take option, a strong disinvestment behavior is observed in several of the non-successful groups. We do not find that heterogeneity in operating funds has a significant negative effect on group success. Consistent with the literature, we find that a higher potential loss rate leads to more groups reaching the threshold.

1.3 Individual contribution to the research papers included in the dissertation

In this section, I describe my contribution to the production of the individual research papers. I follow the CRediT author statement (Contributor Roles Taxonomy) used by the publishers Elsevier and MDPI.[10] For each research paper, I list the number of author points attributable to me, as agreed with the respective co-authors. In addition, I indicate who the corresponding author of the research paper is and list at which conferences and seminars the respective research paper was presented.

For the paper "Editorial favoritism in the field of laboratory experimental economics“, I contributed to the following tasks: conceptualization, methodology, validation, formal analysis, investigation, data curation, writing – original draft, writing – review and editing, visualization, and project administration. I presented the paper[11] at the following conferences and university seminar: Maastricht Behavioral Economic Policy Symposium (M-BEPS); ASSA/AEA 2021 Virtual Annual Meeting; 2019 Economic Science Association (ESA) North American Meeting in Los Angeles (USA); 14th Nordic Conference on Behavioral and Experimental Economics in Kiel; Clausthaler Ökonomisches Oberseminar. In addition, the paper was accepted for the following conferences and events which were cancelled due to the Covid-19 pandemic: Thurgau Experimental Economics Meeting (theem), Kreuzlingen (Switzerland); Innsbruck Winter School on Credence Goods, Incentives and Behavior, Kühtai (Austria); GSBE-ETBC Seminar, Maastricht University (Netherlands). I am the corresponding author of the paper and the co-authors have agreed that a share of 0.7 author points is allocated to me.

For the paper "Is your privacy for sale? An experiment on the willingness to reveal sensitive information", I contributed to the following tasks: conceptualization, methodology, investigation, data curation, writing – original draft preparation, formal analysis, visualization, writing – review and editing, and project administration. I presented the paper at the following conference and university seminar: 2019 Annual Meeting of the German Association for Experimental Economic Research (GfeW e.V.) in Düsseldorf (together with Marco Nilgen); Clausthaler Ökonomisches Oberseminar. Together with Björn Frank, I am the corresponding

[10] See https://www.elsevier.com/authors/policies-and-guidelines/credit-author-statement, https://res.mdpi.com/data/contributor-role-instruction.pdf (both accessed April 09, 2021).

[11] The conference title of the article was “Geographical concentration and editorial favoritism within the field of laboratory experimental economics”.

author of the paper. Since the paper is the result of a one-week PhD course, we have agreed with all 10 authors that each author is allocated a share of 0.1 author points.

I am the single author of the paper "Employer review platforms – do the rating environment and platform design affect the informativeness of reviews? Theory, evidence, and suggestions". Therefore, I carried out all tasks in the production of the paper and have a share of 1 author point.

For the paper "Acceptance of data sharing in smartphone apps from key industries of the digital transformation: A representative population survey for Germany", I contributed to the following tasks: conceptualization, methodology, formal analysis, investigation, data curation, writing – original draft, writing – review and editing, visualization, supervision, and project administration. I presented the paper at the following university seminars: Diginomics Brownbag Seminar at the University of Bremen; Clausthaler Ökonomisches Oberseminar. The paper is currently under review at *Technological Forecasting and Social Change*. I am the corresponding author of the paper and the co-author has agreed that a share of 0.5 author points is allocated to me.

For the paper "Recycling behavior of private households: an empirical investigation of individual preferences in a club good experiment", I contributed to the following tasks: writing – original draft, and writing – review and editing. The corresponding author of the paper is Jacob Wehrle. The co-authors have agreed that a share of 0.16 author points is allocated to me.

For the paper "Combating climate change: Is the option to exploit a public good a barrier for reaching critical thresholds? Experimental evidence", I contributed to the following tasks: conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing – original draft, writing – review and editing, visualization, supervision, and project administration. I presented the paper at the following conferences: 2021 Annual Meeting of the German Association for Experimental Economic Research (GfeW e.V.) in Magdeburg; Virtual 96th Annual Conference of the Western Economic Association International (WEAI). I am the corresponding author of the paper and the co-author has agreed that a share of 0.75 author points is allocated to me.

2. Editorial favoritism in the field of laboratory experimental economics

This chapter has appeared as:

Cloos, J., Greiff, M., & Rusch, H. (2021). *Editorial favoritism in the field of laboratory experimental economics (RM/20/014-revised-)*. Maastricht University, Graduate School of Business and Economics. GSBE Research Memoranda, No. 005.

Abstract: We examine scientific quality and editorial favoritism in the field of experimental economics. We use a novel data set containing all original research papers (N = 569) that exclusively used laboratory experiments for data generation and were published in the *American Economic Review* (AER), *Experimental Economics* (EE), or the *Journal of the European Economic Association* (JEEA) between 1998 and 2018. Several proxies for scientific quality indicate that experiments conducted in Europe are of higher quality than experiments conducted in the US: European experiments rely on larger numbers of participants as well as participants per treatment and receive more citations. For the AER and the JEEA, but not for EE, we find that papers authored by economists with social ties to the editors receive significantly fewer citations in the years following publication. Detailed analyses using a novel dynamic and continuous measure of the co-authorship distance between editors and authors imply that authors at longer distances to editors have to write papers of higher quality in order to get published in the AER and the JEEA. We find no evidence that this 'uphill battle' is associated with geographical distance. (JEL: A11, A14, C90, I23)

Keywords: laboratory experiments, favoritism, methodological standards, network effects

2.1 Introduction

The editorial process is a crucial element in scientific competition. Authors compete for sparse space in journals, ideally leading to only the best manuscripts being accepted for publication. Across journals, editors compete because each editor wants to maximize the quality of her journal by publishing the best papers (Card and DellaVigna, 2020). If decisions about acceptance and rejection are made solely based on the quality of the manuscript, competition ensures that optimal outcomes are achieved. The resulting allocation is efficient in the sense that the best rejected manuscript is of lower quality than the worst accepted manuscript. In judging the quality of a submitted manuscript, editors play a crucial role. However, when editorial decisions are not based on scientific criteria alone but also on others, such as authors' social ties to editors, outcomes might be inefficient and scientific progress might suffer.

Editorial favoritism is a problem, thus, when the quality of published research is negatively affected, i.e., when papers written by an editor's colleagues or former co-authors are published while better papers by other scholars are rejected. High shares of published papers by authors who are connected to at least one of a given journal's editors are often taken as an indication for editorial favoritism. However, this is neither necessary nor sufficient. As pointed out by Brogaard et al. (2014, p. 252), the decisive question is "[…] whether editors use information advantages to improve selection decisions, or whether they bow to conflicts of interest."

Existing studies on effects of author-editor connections (i.e. social ties) use papers published in – usually high ranking general interest – journals (Colussi, 2018; Brogaard et al., 2014; Medoff, 2003; Laband and Piette, 1994a). The studies by Colussi (2018) and Brogaard et al. (2014) show that the number of papers an author publishes in a given journal increases significantly as soon as a close colleague becomes an editor at that journal. Brogaard et al. (2014), Medoff (2003), and Laband and Piette (1994a) show that papers by authors with social ties to editors on average receive significantly more citations in the years following publication. However, the results of Laband and Piette (1994a) also show that more than two thirds of the papers with remarkably few citations were written by authors who had social ties to the editors.

Taken together, the results from previous studies suggest that (1) social ties can improve the chances of getting published because they reduce editors' search costs for high-quality papers, and (2) that, conditional on being accepted for publication papers by authors with social ties to editors receive more citations on average.

Most of the existing literature on editorial favoritism proxies the quality of papers solely *ex post*, i.e. based on the number of citations a paper receives in the years following publication. Then, previous work typically assesses how the number of citations is affected by social ties, authors' reputation, paper length, JEL codes, and journal. In the field of experimental economics, however, several characteristics of an experiment – such as the number of participants, the number of participants per treatment, the number of treatments, and the strength of monetary incentives – constitute *ex ante* proxies for an experiment's quality which typically also affects the respective paper's quality.

While the role of such experimental characteristics has been discussed from a methodological perspective, their role in the in the editorial process has not been studied yet. With this paper, we fill that gap. Our analyses cover all laboratory experiments published between 1998 and 2018 in the *American Economic Review* (AER), arguably one of the top journals for general economics worldwide, *Experimental Economics* (EE), the top field journal, and the *Journal of the European Economic Association* (JEEA), arguably one of the European top journals for general economics.

Our focus on experimental economics allows us to examine (1) whether papers written by authors from different geographical regions differ with respect to several *ex ante* proxies capturing the scientific quality of the experiments, and (2) whether social ties between editors and authors have an effect on the *ex post* quality of papers (measured in citations) when we control for *ex ante* quality proxies.

Our results show substantial differences between Europe and the US, the two regions where more than 82% of the experiments in our dataset were conducted. For all three journals, experiments conducted in Europe have larger average numbers of participants per treatment. In the AER and in EE*,* experiments conducted in Europe also have a larger average total number of participants. We also find that AER and EE papers that include experiments conducted in Europe receive significantly more citations in the years following publication.

For all three journals, our results show that the share of US-affiliated authors of a paper has a negative effect on the number of citations this paper receives. For AER and JEEA papers, our measure for the co-author distance between editors and authors also shows that papers by authors at larger distances to editors receive significantly more citations. Moreover, additional

binary connection measures show that AER and JEEA papers by authors with social ties to editors receive fewer citations in the years following publication.

Our results also indicate that there seems to be no methodological consensus between geographical regions on how to conduct economic experiments – or, equivalently, that there are different 'experimental schools'. Moreover, our results suggest that journal editors treat authors differently, not only with respect to their social ties but also based on authors' affiliations.

Our study provides important insights that editors can use in order to make their selection of research papers more efficient and fairer. Thus, we provide suggestions on how to increase impact factors and simultaneously further scientific progress.

The remainder of this paper is organized as follows. In Section 2.2, we describe the selection criteria for journals and papers, the data set, the *ex ante* quality proxies, and our measures for social ties. In Section 2.3, we compare how quality proxies and social tie measures differ between journals and between geographical regions before we test for editorial favoritism. In Section 2.4, we discuss the results, derive policy implications and conclude.

2.2 Data collection and rationale for variable construction

2.2.1 Selection criteria

Although we are interested in the broad field of experimental economics, we decided to focus exclusively on papers that report results from laboratory experiments. Precisely, we focus on laboratory experiments that generate data in a controlled process using student participants who interact in an artificial environment[12]. In order to ensure the highest possible comparability, we do not consider papers that contain other types of experiments.

For laboratory experiments, objectively measurable proxies for an experiment's quality exist. Proxies for an experiment's quality are the *total number of participants*, the *number of treatments*, the n*umber of participants per treatment*, and the strength of monetary incentives (henceforth, *strength of incentives*). When conducting a laboratory experiment, the

[12] Our data contain papers where the majority of the participants were students. Papers where a large share of participants belonged to special groups (such as job professionals or caste members etc.) are not included in the data set.

experimenter has full control over these characteristics. This is generally not the case for field experiments.

We chose the AER and the JEEA because they are general interest journals and the official journals of the *American Economic Association* and the *European Economic Association*, respectively. Moreover, the AER publishes the largest number of experimental papers among the top-5 economics journals (Nikiforakis and Slonim, 2019).We did not include any comments or papers from AER's *Papers and Proceedings* in the dataset. EE is the top field journal, and papers published there consist predominantly of laboratory experiments. We chose 1998 as the starting year for our analysis because EE was founded that year. EE is the specialized field journal of the *Economic Science Association*.

2.2.2 Data set

Our data set contains 569 papers published between 1998 and 2018 (407 from EE, 121 from the AER, and 41 from the JEEA)[13].

For each paper in our data set, we collected data on the:

- authors and their affiliations at the time of publication,
- number of female authors,
- number of citations received by each author in the five years before publication,
- total number of pages,
- total number of references, and
- paper's number of citations as of mid-November, 2020.

For all papers, we extracted data on the characteristics of the laboratory experiments reported in the papers. For each laboratory experiment we collected data on the:

- total number of participants,
- number of treatments,
- duration of the experiment,
- average earnings per participant,
- year in which the experiment was conducted, and
- place (laboratory) where the experiment was conducted.

[13] A full list of the included papers can be found in the Appendix (section 2.5.2).

In case a paper reported results from more than one laboratory experiment, we computed average values for the first four characteristics. Since not all papers report data on all characteristics, our data set contains some papers for which some or all of the experiment's characteristics are missing. In case of missing values, we tried to obtain the data from working paper versions or directly from the authors, which was successful in some cases. Citation data was obtained from *Google Scholar* (GS).

To identify connections between authors and editors, we collected the names and affiliations of all editors and co-editors for each of the three journals, as well as the years in which they served as editors or co-editors. Due to the time lag between submission and publication, we consider economists who served as editor or co-editor between 1996 and 2016.

We decided to focus on editors and co-editors and not on referees for three reasons. First, editors have the power to overrule referees' recommendations. Second, editors tend to select referees to whom they have easy access, implying that social ties between author and editor and between author and referee are correlated (Hamermesh, 1994). And third, data on referees is hard to obtain (an exception is Card et al., 2020).

2.2.3 Proxies for the quality of laboratory experiments

Our quality proxies can be divided into *ex ante* and *ex post* proxies. *Ex ante* proxies are available before a paper is published and can thus be used by editors when deciding whether to reject or accept a paper. *Ex post* proxies measure the quality of a paper after it has been published.

The first two *ex ante* proxies are the *total number of participants* and the *number of participants per treatment*. Both are related to the experiment's statistical power. Experiments with higher power are more likely to detect small effects, tend to generate fewer false positives and, hence, produce results that are more likely to replicate. This in turn could increase confidence in the experiment's results. When designing their experiments, experimental economists often use simple rules of thumb, like those explained in List et al. (2011). However, Zhang and Ortmann (2013) show that many experiments are underpowered and that rules of thumb provide insufficient information about the power of a study. For detailed discussions of power analysis in experimental economics see Vasilaky and Brock (2020), Czibor et al. (2019), Ioannidis et al. (2017), Bellemare et al. (2016, 2014), and Zhang and Ortmann (2013).

For a given number of treatments, a larger *total number of participants* yields more observations which increases statistical power if statistical testing is carried out at the participant-level. The

number of participants per treatment is a more suitable proxy for the experiment's power if an experiment contains a large number of treatments and if the statistical analysis involves a pairwise comparison of treatments. Arguably, another possibility is to consider the number of independent observations as a proxy for quality. However, what constitutes an independent observation is not clearly defined. Thus, we decided not to follow this approach (also see chapter 3 in Svorenčík and Maas, 2016).

The third *ex ante* proxy is the *number of treatments.* Ceteris paribus, a larger *number of treatments* increases quality because it allows to test more research hypotheses and/or to rule out more alternative explanations. Both increase quality, because tests of more hypotheses increase a paper's scientific contribution and ruling out more alternative hypotheses yields more credibility to the authors' claims.

The fourth proxy is the *strength of incentives.* According to the methodological literature on laboratory experiments, monetary incentives are related to the quality of an experiment because they tend to mitigate experimenter demand effects. Additionally, participants exert more effort in judgement and decision tasks, which reduces variance and the resulting data is less noisy (Bardsley et al., 2010; Camerer and Hogarth, 1999). The use of monetary incentives is an established methodological standard in economics (see Hertwig and Ortmann, 2001, p. 390) and all experiments in our data set use monetary incentives. We operationalized the *strength of incentives* by dividing a participant's average earnings (including the show-up fee) by the duration of the experiment (in minutes) and converting the result into real 2015 US-Dollars.

We focus on only one *ex post* proxy, namely citations[14]. This does not imply that citations reflect a paper's true quality. Rather, we decided to use the *number of citations* because data on citations is widely available and is heavily used (at least as a yardstick) to allocate positions and resources (Card et al., 2020; Card and DellaVigna, 2020; Hamermesh, 2018; Moed, 2006; Laband and Piette, 1994b). Moreover, assuming that editors seek to maximize their journals' impact factor, they have an incentive to accept papers that they expect to receive a high number of citations.

[14] Other *ex post* proxies are being reprinted in anthologies, being discussed in canonical textbooks, or the amount of media-coverage.

2.2.4 Measures for social ties between editors and authors

In the following, we introduce four different ways to measure social ties: the *share of US authors*, *minDist* (our novel measure based on co-authorship data), *colleague connection*, and *co-author connection*.

Between 1996 and 2016, the AER had 32 different editors, of which 31 (96.88%) had their affiliation at US universities and one (3.12%) in the UK. EE had eight different editors, of which five (62.5%) had their affiliation at a US university and three (37.5%) at a European university. The JEEA, founded in 2003, had 18 different editors, of which eight (44.44%) had their affiliation at a US university and ten (55.56%) at a European university. Given the high number of US-based editors, especially on the editorial board of the AER, we use the *share of US authors* for each paper as a first measure of social ties between editors and authors. Here we implicitly assume that a social tie is more likely to exist if editor and author have their affiliation within the same country (the US in case of the AER).

All previous studies on editorial favoritism use binary measures for social ties between editors and authors. For example, a social tie is assumed if editor and author had their affiliation at the same institution in the period of publication. We introduce a novel measure, which is based on co-authorship data. Specifically, we compiled a large dynamic network structure allowing us to compute all co-author distances between the 58 editors and the 931 authors in our data set at the time of publication of the respective paper. The calculation is based on approximately 450,000 unique papers published between 1950 and 2020 in 1,434 economics journals written by 268,000 authors. For each paper, we compute the distances between each author of the paper and all editors of the respective journal. Our novel measure, *minDist*, then is the minimum distance between all authors of a paper and all editors of the journal for the corresponding publication period. In line with Brogaard et al. (2014) we look at all editors of the respective journal two years before the paper was published. If no co-author connections could be identified between a single author (or between a group of authors) and the editors, we did not assign a value to *minDist*. These observations (N = 10) are not considered in the following results on *minDist*. Among the 586 papers for which we were able to calculate *minDist*, the values range from 0 to 9 (mean = 4.53, sd = 1.52, median & mode = 5).

In addition to *minDist*, we use two commonly used binary measures of social ties. We assume a *colleague connection* if the author has the same affiliation at the time of publication like one of the editors of the respective journal two years before publication. A *co-author connection*

exists if an author of the paper is also the editor of the journal (i.e. *minDist* = 0), or if the author of the paper is a former co-author of one of the editors (*minDist* = 1).

2.3 Results

2.3.1 How do quality proxies and social ties differ between journals?

The first five rows of Table 2.1 show mean values for quality proxies and social ties for the three journals and *p*-values from pairwise comparisons across journals. The *total number of participants,* and the *number of treatments* are significantly higher for the AER and the JEEA compared to EE (two-sided Mann-Whitney tests, all *p*-values < 0.05). The variable *strength of incentives* is significantly higher for the AER compared to EE (*p*-value < 0.01). The *strength of incentives* could only be computed for 54% of AER papers, 72% of EE papers and 66% of JEEA papers. This is because, even in these prestigious journals, many papers do not report the number of participants, average earnings per participant, or the duration of the experiment. The means of citations per year do significantly differ between all three journals and are highest for the AER, followed by the JEEA.

Rows six to eight show mean values for social ties between authors and editors. In 10.74% of the AER papers and 17.07% of the JEEA papers, at least one author of the paper has a *colleague-connection* to one of the editors. These values are significantly higher than for EE (two-sided Fisher's exact tests, both *p*-values < 0.01), where a *colleague-connection* is found for only 3.69% of the papers. The average values for *minDist* are also significantly lower for the AER and JEEA than for EE (two-sided Mann-Whitney tests, both *p*-values < 0.001).

Row #		AER (N=121)	EE (N=407)	JEEA (N=41)	AER vs. EE	AER vs. JEEA	EE vs. JEEA
				Ex ante quality proxies			
1	# participants	273.78 (N=114)	194.83 (N=404)	284.58 (N=40)	$p<0.001$	$p=0.396$	$p<0.001$
2	# treatments	4.70 (N=120)	3.91 (N=404)	4.74 (N=35)	$p=0.011$	$p=0.538$	$p=0.023$
3	# participants per treatment	71.46 (N=113)	55.22 (N=402)	61.82 (N=34)	$p=0.266$	$p=0.581$	$p=0.202$
4	strength of incentives	0.38 (N=70)	0.31 (N=293)	0.33 (N=27)	$p<0.001$	$p=0.159$	$p=0.482$
				Ex post quality proxy			
5	citations per year	33.48 (N=118)	6.50 (N=342)	20.76 (N=31)	$p<0.001$	$p=0.015$	$p<0.001$
				Social Ties			
6	colleague connection in %	10.74	3.69	17.07	$p=0.005$	$p=0.284$	$p=0.002$
7	co-author connection in %	4.96	4.42	0.00	$p=0.806$	$p=0.338$	$p=0.392$
8	any connection in %	14.05	7.37	17.07	$p=0.029$	$p=0.619$	$p=0.065$
9	minDist	3.98	4.76	3.93	$p<0.001$	$p=0.415$	$p<0.001$

Table 2.1: Quality proxies and social tie measures by journal.

Notes. Mean values and number of observations. For rows 1, 2, 3, 4, 5, and 9, the three right columns show the exact p-values for two-sided Mann-Whitney tests. For rows 6, 7, and 8, they show p-values for two-sided exact Fisher tests. For the calculation of citations per year we only include papers that were published before 2017.

2.3.2 How do quality proxies and social ties differ between the US and Europe?

In Table 2.2, we look only at laboratory experiments conducted in the US or Europe. Slightly more than 82% of the laboratory experiments in our dataset were conducted in these two regions.

With respect to the *total number of participants* or the *number of participants per treatment*, there appear to be systematic differences between experiments conducted in both regions. Except for the JEEA, experiments conducted in Europe involve a significantly larger *total number of participants*. For all journals, European experiments have a significantly larger *number of participants per treatment*. With regard to the *number of treatments* and the *strength of incentives*, there are no significant differences between North-American and European experiments.

For the AER and EE we observe that laboratory experiments conducted in Europe receive significantly more *citations per year* compared to laboratory experiments conducted in the US.

For colleague and co-author connections, we do not find significant differences between the US and Europe. However, we observe that for AER and EE papers the social tie measure *minDist* has significantly lower values for experiments conducted in the US.

Row #		**AER (N=105)**		**EE (N=324)**		**JEEA (N=38)**	
				Ex ante quality proxies			
		US	**Europe**	**US**	**Europe**	**US**	**Europe**
1	# participants	234.09 (N=65)	380.56 (N=34)	167.97 (N=154)	205.60 (N=167)	243.57 (N=14)	295.83 (N=23)
		p=0.002		p=0.010		p=0.957	
2	# treatments	4.61 (N=69)	5.00 (N=35)	3.86 (N=155)	3.81 (N=166)	5.00 (N=12)	4.48 (N=21)
		p=0.674		p=0.550		p=0.844	
3	# part. per treatment	55.24 (N=64)	113.70 (N=34)	49.68 (N=153)	59.90 (N=166)	37.90 (N=11)	73.53 (N=21)
		p<0.001		p=0.007		p=0.012	
4	strength of incentives	0.35 (N=37)	0.45 (N=28)	0.30 (N=108)	0.33 (N=130)	0.36 (N=7)	0.31 (N=18)
		p=0.050		p=0.082		p=0.745	
				Ex post quality proxy			
5	citations per year	30.83 (N=69)	44.85 (N=33)	5.47 (N=136)	7.62 (N=143)	20.52 (N=12)	21.30 (N=17)
		p=0.047		p=0.006		p=0.711	
				Social Ties			
6	colleague connection in %	15.71	5.71	3.82	3.59	13.33	21.74
		p=0.211		p=1.000		p=0.681	
7	co-author connection in %	2.90	8.57	5.77	2.50	0.00	0.00
		p=0.332		p=0.166		p=1.000	
8	any connection	17.14	11.43	8.92	5.39	13.33	21.74
		p=0.570		p=0.280		p=0.681	
9	minDist	3.75	4.31	4.31	5.26	4.14	3.87
		p=0.003		p<0.001		p=0.176	

Table 2.2: Quality proxies and social tie measures by journal and region.

Notes. Mean values and number of observations. The rows below rows 1, 2, 3, 4, 5, and 9, show exact p-values for two-sided Mann-Whitney tests. The rows below rows 6, 7, and 8, show p-values for two-sided exact Fisher tests. For the calculation of citations per year we only include papers that were published before 2017.

2.3.3 Is there evidence for editorial favoritism in laboratory experimental economics?

In the previous subsection, we saw that experiments conducted in Europe receive more citations than experiments conducted in the US. In this section, we investigate whether the differences in citations are mainly driven by differences in *ex ante* quality proxies or by social ties between authors and editors. Our *ex ante* proxies are scientific criteria because they are derived from methodological considerations (see section 2.2.3). Hence, we would expect them to influence the quality of a paper, which we proxy by the number of citations. Social ties, however, are

non-scientific criteria, and if they have a negative effect on a paper's quality, this would constitute evidence of quality reducing editorial favoritism.

We start by looking at the relation between the inverse hyperbolic sine (asinh)[15] of *citations per year* and *minDist* in Figure 2.1. For AER and JEEA papers, we see that there is a positive correlation between *citations per year* and *minDist*. Thus, Figure 2.1 shows how papers by authors at larger distances to editors receive a higher number of citations than papers written by authors with closer connections to the editors.

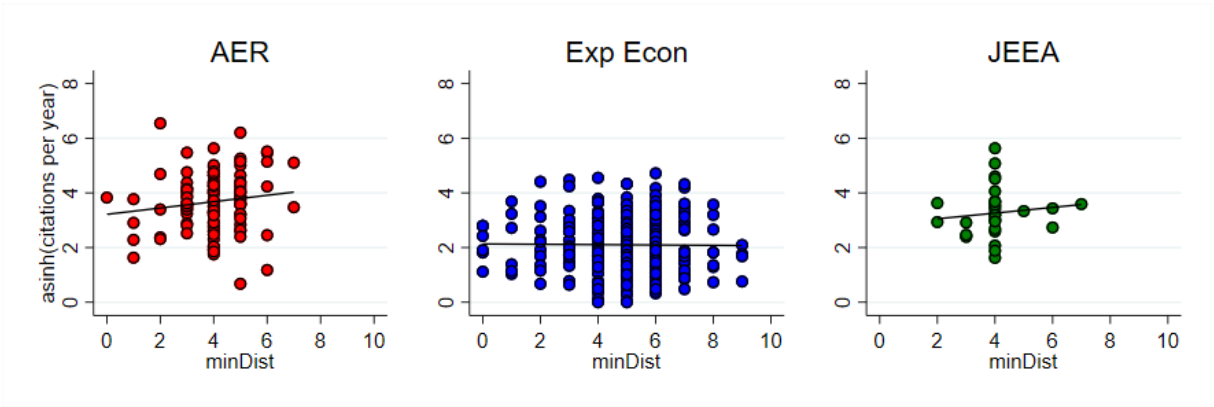


Figure 2.1: Scatterplots of minDist and asinh(citations per year) by journal.

Figure 2.1 is based on raw data, that is, we do not control for differences in *ex ante* quality characteristics and other factors, which might affect the number of citations a paper receives. Following Brogaard et al. (2014), Medoff (2003), and Laband and Piette (1994a) we estimate the following regression model:

$$\text{asinh}(C_i) = \beta_0 + \beta_1 AER_i + \beta_2 JEEA_i + \beta_3 ST_i + \beta_4 (ST_i * AER_i) + \beta_5 (ST_i * JEEA_i) + \beta_6 age_i + \gamma E_i + \delta P_i + \theta A_i + u_i \qquad (1)$$

The dependent variable $\text{asinh}(C_i)$ is the inverse hyperbolic sine of the total number of citations a paper has received by mid-November, 2020. To ensure that each paper had sufficient time to receive citations, we only include papers that were published before 2017.

Our measure of social ties is ST_i. Dependent on the regression model it indicates either the relative share of US-affiliated authors, *minDist*, a *colleague connection*, a *co-author connection* or a combination of colleague- and co-author connection (*any-connection*) for paper i. Because

[15] In the following, we use the asinh transformation to account for 0 citations. This practice is in line with previous studies Card and DellaVigna (2020), Card et al. (2020), and Hengel and Moon (2020).

the effect of social ties could differ between journals, we include interaction terms for the respective social tie measure and indicator variables for journals.

We control for each paper's age, as well as for several other characteristics that might affect the number of citations. E_i is a vector containing the *ex ante* quality proxies of the experiments reported in paper i. By including indicator variables for journals, we control for journal fixed effects. The vector P_i contains controls for number of (EE-equivalent) pages, number of references, and the JEL-classification. A_i is a vector containing author related controls for gender (share of female authors), reputation (average of citations received in the five years prior publication), and number of authors.

Table 2.3 shows estimation results for different OLS regression models with standard errors clustered by journal (for full results see Appendix Tables 2.5 and 2.6). In Panel A, we examine the influence of the *share of US authors* (models 1 to 3) and *colleague connections* (models 4 to 6). The results show that, for all three journals, the *share of US authors* has a negative effect on the number of citations. In models 1 and 2, this effect is even stronger for AER and JEEA papers. For EE papers, the results of models 4 to 6 indicate that *colleague connections* have a significant positive effect on citations. In contrast, for AER and JEEA papers we observe significantly negative effects of *colleague connections*.

In the lower half of Table 2.3 (Panel B), we examine the effect of the minimum co-authorship distance (*minDist*) and a direct authorship or co-authorship (*co-author connection*, *minDist* < 2) on the number of citations. We observe significant positive effects for the interaction terms of *minDist* with AER or JEEA. This implies that papers by authors at longer distances receive significantly more citations than papers by authors with closer connections to the editors. For the AER, this effect is also apparent when we consider direct *co-author connections*. AER papers written by the editors themselves or by their co-authors have a significantly lower number of citations. When we use the raw number of citations as dependent variable (see Appendix Table 2.11), the results show that a one point increase in *minDist* for AER papers is associated with an increase between 116 and 133 citations.

The variable *colleague connection* is significantly positive in all three models in Panel A. The variable *co-author connection* is also significantly positive in model 4 of panel B. This could imply that EE editors use their professional connections to identify high quality papers for their journal.

Among the *ex ante* quality proxies, only the *number of treatments* and the *strength of incentives* have significant positive effects on citations in several models (see Appendix Tables 2.5 to 2.12).

Panel A						
social ties: share of US authors and colleague-connected papers						
	dependent variable: asinh(GS citations)					
	(1)	(2)	(3)	(4)	(5)	(6)
share US authors	-0.16***	-0.12***	-0.13***			
	(0.01)	(0.01)	(0.01)			
share US auth. * AER	-0.41***	-0.27**	-0.11			
	(0.00)	(0.04)	(0.12)			
share US auth. * JEEA	-0.14***	-0.33**	0.04			
	(0.00)	(0.03)	(0.22)			
colleague con.				0.40***	0.36***	0.14**
				(0.00)	(0.01)	(0.03)
colleague con. * AER				-0.64***	-0.51***	-0.25
				(0.01)	(0.02)	(0.14)
colleague con. * JEEA				-0.69***	-0.63***	-1.06***
				(0.01)	(0.04)	(0.08)
E	No	Yes	Yes	No	Yes	Yes
P	No	No	Yes	No	No	Yes
A	No	No	Yes	No	No	Yes
N	329	327	318	329	327	318
Adjusted R^2	0.41	0.41	0.49	0.40	0.41	0.49
Panel B						
social ties: minDist and coauthor-connected papers						
	dependent variable: asinh(GS citations)					
	(1)	(2)	(3)	(4)	(5)	(6)
minDist	-0.01***	-0.01	-0.01			
	(0.00)	(0.00)	(0.00)			
minDist * AER	0.33***	0.33***	0.22*			
	(0.00)	(0.01)	(0.05)			
minDist * JEEA	0.14***	0.15***	0.24***			
	(0.00)	(0.01)	(0.02)			
co-author connection				0.06***	-0.01	0.14
				(0.01)	(0.03)	(0.05)
co-aut. con. * AER				-0.86***	-0.92***	-0.87**
				(0.01)	(0.06)	(0.20)
co-aut. con. * JEEA				n.a.	n.a.	n.a.
E	No	Yes	Yes	No	Yes	Yes
P	No	No	Yes	No	No	Yes
A	No	No	Yes	No	No	Yes
N	321	319	310	321	319	310
Adjusted R^2	0.41	0.42	0.49	0.40	0.41	0.48

Table 2.3: Impact of social ties on ex post paper quality.

Notes. OLS regressions with standard errors clustered by journal in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each regression model includes controls for journal and years since publication.

At least for the *share of US authors*, the results of Table 2.3 may at least partially be driven by network effects. In 2018, more than half (57.84%) of all experimental economists on the RePEc

list (N = 1831) of authors in experimental economics had a European affiliation, followed by economists with an affiliation in North America (30.20%), Asia (5.90%), Australia or New Zealand (3.82%), South America (1.86%), and Africa (0.38%). In the following, we therefore examine whether our results change when we consider solely US papers.

Table 2.4 shows the results for regressions exclusively including papers with experiments conducted in the US and where the share of US affiliated authors is at least 0.5 (for full results see Appendix Table 2.7). In models 1 and 2, the variable *minDist* again has a significant positive effect on the number of citations for AER and JEEA papers. Due to the small number of observations, we combined the variables *colleague connection* and *co-author connection* into the variable *any connection*. In models 4 and 5, we observe that AER papers by authors with connections to the editors receive significantly fewer citations.

US papers						
social ties: minDist (models 1-3); any connection (colleague connection or co-author connection)						
	dependent variable: asinh(GS citations)					
	(1)	(2)	(3)	(4)	(5)	(6)
minDist	-0.01	-0.00	-0.03*			
	(0.00)	(0.02)	(0.01)			
minDist * AER	0.21***	0.25**	0.20			
	(0.00)	(0.05)	(0.11)			
minDist * JEEA	0.84***	0.63*	0.11			
	(0.02)	(0.22)	(0.04)			
any-connection				0.22**	0.10*	0.07
				(0.03)	(0.02)	(0.09)
any-connection * AER				-0.45***	-0.23*	-0.03
				(0.04)	(0.06)	(0.11)
any-connection * JEEA				n.a.	n.a.	n.a.
E	No	Yes	Yes	No	Yes	Yes
P	No	No	Yes	No	No	Yes
A	No	No	Yes	No	No	Yes
N	113	112	111	116	115	114
Adjusted R^2	0.34	0.36	0.42	0.36	0.37	0.44

Table 2.4: Impact of social ties on ex post paper quality (only US papers).

Notes. OLS regressions with standard errors clustered by journal in parentheses: * p < 0.10, ** p < 0.05, *** p < 0.01. Each regression model includes controls for journal and years since publication.

Regarding the *ex ante* quality proxies (suppressed in Tables 2.3 and 2.4), positive signs are observed for the *number of treatments* and the *strength of incentives*. These effects, however, are not always significant. We find no effect for the share of female authors. Our results from Tables 2.3 and 2.4 are largely robust when we use citations from *Web of Science* (WoS) as the dependent variable (see Appendix Tables 2.8 and 2.9) or when we use alternative specifications of the dependent variable (see Appendix Tables 2.10, 2.11, and 2.12).

2.4 Discussion and conclusion

In this paper we use a novel data set containing data on all laboratory experiments published in the *AER*, EE and the *JEEA* from 1998 to 2018. We identify and compare several objectively measurable quality proxies between journals and between the US and Europe and test for editorial favoritism. Our study adds several aspects to the existing literature on editorial favoritism.

First, focusing on one particular field allows us to look not only at citations as an *ex post* quality proxy, but also to analyze how various *ex ante* quality proxies which are available only for laboratory experiments (total number of participants, participants per treatment, number of treatments, strength of incentives) differ between journals and regions (which may follow different methodological traditions). Second, we conduct an empirical analysis on the role of the above-mentioned quality characteristics and social ties between authors and editors. In this context we introduce *minDist*, a new measure for social ties which is based on the co-authorship distance between authors and editors. In contrast to existing binary measures of social ties, *minDist* is more fine-grained since it allows to capture indirect co-author connections between editors and authors. In our calculation of *minDist* we also considered significantly more journals than previous studies used to identify co-author connections. Thus, the use of *minDist* allows us to investigate the effects of (co-author based) social ties also cases like the JEEA where no direct co-author connections exist between editors and authors (see Table 2.3). Nevertheless, *minDist* also has limitations. For one, it is not straightforward how to interpret large *minDist* values: is the co-author of your co-author's co-author's co-author still someone you would favor? For the other, we considered all co-authorship connections in the calculation of *minDist* and did not impose any restriction on the time span. As a result, several *minDist* values may be based on co-authorship connections that lie relatively far in the past. However, a restriction of the time span would require a definition of when a connection should be classified as expired. Since such a definition would necessarily be arbitrary, we did not impose a restriction.

Our results show that various *ex ante* quality proxies are higher for AER and JEEA papers than for EE papers. This also reflects the higher number of citations per year that papers published in the former two journals received. Comparing experiments conducted in the US and Europe, there are no differences regarding the *number of treatments* and no or only small differences regarding the *strength of incentives*. However, European experiments published in the AER and EE rely on a significantly larger *total number of participants* and in all three journals on a

significantly larger *number of participants per treatment*. This difference is most pronounced for the AER where, on average, experiments conducted in the US have 55 *participants per treatment* while experiments conducted in Europe have 113 *participants per treatment*.

The differences in the *total number of participants* and *participants per treatment* could reflect different methodological standards. However, they could also indicate substantial barriers to entry for European economists, especially for the AER, where US-dominance is more pronounced in the composition of the editorial board. Our results for the *total number of participants* and the *strength of incentives* imply that experimental economists from Europe pay a higher price for publishing their papers, compared to experimental economists from the US. However, our data cannot answer if this is due to barriers of entry or due to economists in Europe having easier access to research funding compared to their colleagues in the US[16].

Different measures of social ties show that there are closer social ties between authors and editors for AER and JEEA papers than there are for EE papers. A potential explanation for this result is that the number of EE editors during the respective period was significantly smaller than the number of editors in the other two journals.

Our investigation of editorial favoritism shows that the share of US authors in the top general interest journals AER and JEEA has a significant negative effect on the number of citations a paper received. For AER and JEEA papers, several measures of social ties indicate that papers by authors with social ties or closer connections to the editors are of lower quality than papers without these social ties.

These differences can be interpreted as indirect evidence of discrimination, indicating that a paper from an author with social ties is accepted for publication despite a higher-quality paper from an author without social ties being available. Of course, this only holds true if high-quality papers by authors without social ties to the editorial board were submitted and rejected.

[16] We thank Gary Charness for drawing our attention to this aspect. Unfortunately, there are no data on the amount of research funding for experimental economists in different regions. However, we consider it unlikely that there is a causal relationship between the availability of research funding and the average *total number of participants*. Even if research funds were more readily available in Europe, they could be used, for example, to conduct more (single) projects.

However, given the high rejection rates of the AER and the JEEA, we think that this is quite plausible[17].

Note, however, that our results do not imply that editors consciously discriminate against authors from outside the US or against authors at longer co-authorship distances to themselves. Rather, it is possible that editors systematically fail to predict the number of citations a paper will receive in the years following its publication (i.e., editors might wrongly predict that papers authored by US-based and/or connected authors receive more citations than they actually do). Nonetheless, if citations can be taken as a proxy for quality, this will slow down scientific progress, and editors will fail to maximize their journal's impact factor.

Except for the share of US authors, we find no negative or significantly positive effects of social ties for EE papers. This might be explained by the fact that EE is a field journal. EE editors usually have an in-depth expertise in experimental economics and may thus be particularly well qualified to identify high quality experimental papers. Editors of the AER and the JEEA do not necessarily have expertise in experimental economics. While experimental papers are certainly reviewed by referees with an expertise in experimental economics also in these general interest journals, review processes are only initiated if the papers have not been desk-rejected immediately. In order to avoid that potentially promising high quality papers are mistakenly desk-rejected, thus, editors whose expertise only partly matches the method used in a submitted paper could intensify their efforts to obtain second opinions from colleagues who are more competent with regard to this method.

We hope that our results raise an awareness of potential biases in the editorial process among editorial boards and that these biases will thereby be reduced in the longer term. Journals could facilitate further research on editorial favoritism (and discriminating editorial practices in general) by regularly publishing not only numbers on submissions, desk-rejections, revise and resubmit decisions, etc., but also on which countries or institutions the submissions came from. In this context, it would also be useful to report the gender distribution of authors for submissions. By comparing published papers with all submissions, future studies could investigate whether authors from certain countries, from certain institutions, or of a certain gender are over- or under-represented in the published papers. Moreover, in their selection of

[17] In 2019, for example, only 6.95% of all submissions (N = 1,927) were accepted at the AER (Duflo 2020). The JEEA received 1,252 submissions between Q4 2019 and Q3 2020 and desk-rejected 60% of these papers (Source: Tweet by Imran Rasul on January 6, 2021).

new editors and associate editors, editorial board members could pay explicit attention not only to gender fairness but also to geographical diversity (Palser et al., 2021; Angus et al., 2020).

Acknowledgements

We thank Max Albert, Jeanette Brosig-Koch, Gary Charness, Björn Frank, Daniel Hamermesh, Sebastian Kruegel, Sherry Xin Li, Sarah Necker, Hans Joachim Voth, the participants of the ASSA/AEA 2021 Virtual Annual Meeting, the participants of the 2020 Virtual Annual Meeting of the German Economic Association (Verein für Socialpolitik), the participants of the 2019 Economic Science Association (ESA) North American Meeting in Los Angeles, the participants of the 14th Nordic Conference on Behavioral and Experimental Economics in Kiel, the participants of the 2017 Annual Meeting of the German Association for Experimental Economic Research (GfeW e.V.) in Kassel, and seminar participants at the University of Portsmouth, University of Passau, HSU Hamburg and the University of Marburg for feedback and comments. Simone Koch provided excellent research assistance.

2.5 Appendix

2.5.1 Tables

	dependent variable: asinh(GS citations)					
	(1)	(2)	(3)	(4)	(5)	(6)
AER	1.87***	1.72***	1.71***	1.66***	1.57***	1.63***
	(0.00)	(0.05)	(0.07)	(0.01)	(0.02)	(0.09)
JEEA	1.18***	1.22***	0.95**	1.21***	1.15***	1.11***
	(0.00)	(0.03)	(0.14)	(0.00)	(0.00)	(0.02)
share US auth.	-0.16***	-0.12***	-0.13***			
	(0.01)	(0.01)	(0.01)			
share US auth. * AER	-0.41***	-0.27**	-0.11			
	(0.00)	(0.04)	(0.12)			
share US auth. * JEEA	-0.14***	-0.33**	0.04			
	(0.00)	(0.03)	(0.22)			
colleague con.				0.40***	0.36***	0.14**
				(0.00)	(0.01)	(0.03)
colleague con. * AER				-0.64***	-0.51***	-0.25
				(0.01)	(0.02)	(0.14)
colleague con. * JEEA				-0.69***	-0.63***	-1.06***
				(0.01)	(0.04)	(0.08)
age	0.09***	0.09***	0.10***	0.08***	0.09***	0.10***
	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)
part. per treatment		0.00	0.00		0.00	0.00
		(0.00)	(0.00)		(0.00)	(0.00)
treatments		0.04**	0.04		0.04***	0.04
		(0.01)	(0.02)		(0.00)	(0.02)
strength of incentives		0.57*	0.47		0.61*	0.49
		(0.14)	(0.21)		(0.20)	(0.25)
pages (EE-equivalent)			-0.02			-0.02
			(0.02)			(0.02)
references			0.01			0.02
			(0.01)			(0.01)
JEL-micro			-0.08			-0.07
			(0.17)			(0.17)
JEL-theory			-0.38			-0.40*
			(0.15)			(0.12)
JEL-labor			-0.14			-0.11
			(0.31)			(0.28)
JEL-econometrics			-0.81**			-0.78**
			(0.15)			(0.15)
JEL-industrial			-1.51**			-1.58***
			(0.18)			(0.15)
JEL-international			-4.72***			-4.65***
			(0.44)			(0.39)
JEL-finance			-1.08***			-1.03***
			(0.08)			(0.07)
JEL-public			0.94*			0.94
			(0.29)			(0.33)
JEL-health			-1.61			-1.65
			(0.65)			(0.58)
JEL-development			-0.59			-0.46
			(0.38)			(0.30)
JEL-other			0.08			0.02
			(0.40)			(0.39)
share fem. auth.			0.10			0.08
			(0.19)			(0.19)
reputation			0.00			0.00**
			(0.00)			(0.00)
number of auth.			-0.05			-0.04
			(0.05)			(0.05)
Constant	3.40***	3.01***	3.05***	3.38***	2.94***	2.98***
	(0.03)	(0.04)	(0.11)	(0.04)	(0.03)	(0.12)
N	329	327	318	329	327	318
Adjusted R^2	0.41	0.41	0.49	0.40	0.41	0.49

Table 2.5: Impact of social ties on paper quality (full results I).

Notes. OLS regressions with standard errors clustered by journal in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	dependent variable: asinh(GS citations)					
	(1)	(2)	(3)	(4)	(5)	(6)
AER	0.28***	0.16**	0.71	1.67***	1.58***	1.66***
	(0.00)	(0.03)	(0.29)	(0.01)	(0.02)	(0.08)
JEEA	0.54***	0.44**	-0.12	1.10***	1.06***	0.91***
	(0.02)	(0.07)	(0.10)	(0.00)	(0.01)	(0.04)
minDist	-0.01***	-0.01	-0.01			
	(0.00)	(0.00)	(0.00)			
minDist * AER	0.33***	0.33***	0.22*			
	(0.00)	(0.01)	(0.05)			
minDist * JEEA	0.14***	0.15***	0.24***			
	(0.00)	(0.01)	(0.02)			
co-aut. con.				0.06***	-0.01	0.14
				(0.01)	(0.03)	(0.05)
co-aut. con. * AER				-0.86***	-0.92***	-0.87**
				(0.01)	(0.06)	(0.20)
co-aut. con. * JEEA				n.a.	n.a.	n.a.
age	0.08***	0.08**	0.10**	0.08***	0.08***	0.10**
	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)
part. per treatment		0.00	0.00		0.00	0.00
		(0.00)	(0.00)		(0.00)	(0.00)
treatments		0.04**	0.04		0.04***	0.04
		(0.00)	(0.02)		(0.00)	(0.02)
strength of incentives		0.74	0.54		0.70	0.54
		(0.32)	(0.34)		(0.28)	(0.34)
pages (EE-equivalent)			-0.02			-0.02
			(0.02)			(0.02)
references			0.01			0.02
			(0.01)			(0.01)
JEL-micro			-0.08			-0.09
			(0.19)			(0.19)
JEL-theory			-0.40			-0.42
			(0.17)			(0.16)
JEL-labor			-0.34*			-0.37*
			(0.11)			(0.12)
JEL-econometrics			-0.79**			-0.81**
			(0.17)			(0.16)
JEL-industrial			-1.47**			-1.53***
			(0.15)			(0.14)
JEL-international			-5.62***			-5.87***
			(0.25)			(0.05)
JEL-finance			-1.11***			-1.08***
			(0.03)			(0.03)
JEL-public			0.89*			0.99
			(0.27)			(0.36)
JEL-health			-1.58*			-1.72
			(0.52)			(0.61)
JEL-development			-0.69			-0.62
			(0.52)			(0.40)
JEL-other			0.02			0.07
			(0.33)			(0.38)
share fem. auth.			0.06			0.00
			(0.26)			(0.18)
reputation			0.00**			0.00
			(0.00)			(0.00)
number of auth.			-0.02			-0.03
			(0.01)			(0.03)
Constant	3.47***	3.00***	2.99***	3.40***	2.94***	2.95***
	(0.06)	(0.03)	(0.14)	(0.04)	(0.02)	(0.16)
N	321	319	310	321	319	310
Adjusted R^2	0.41	0.42	0.49	0.40	0.41	0.48

Table 2.6: Impact of social ties on paper quality (full results II).

Notes. OLS regressions with standard errors clustered by journal in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	dependent variable: asinh(GS citations)					
	(1)	(2)	(3)	(4)	(5)	(6)
AER	0.68***	0.43	0.70	1.57***	1.44***	1.49***
	(0.01)	(0.15)	(0.47)	(0.01)	(0.05)	(0.06)
JEEA	-2.83***	-2.11	0.25	0.90***	0.80**	0.94***
	(0.08)	(0.75)	(0.42)	(0.01)	(0.12)	(0.06)
minDist	-0.01	-0.00	-0.03*			
	(0.00)	(0.02)	(0.01)			
minDist * AER	0.21***	0.25**	0.20			
	(0.00)	(0.05)	(0.11)			
minDist * JEEA	0.84***	0.63*	0.10			
	(0.02)	(0.22)	(0.05)			
any-connection				0.22**	0.10*	0.07
				(0.03)	(0.02)	(0.09)
any-connection * AER				-0.45***	-0.23*	-0.03
				(0.04)	(0.06)	(0.11)
any-connection * JEEA				n.a.	n.a.	n.a.
age	0.04**	0.05**	0.06***	0.04*	0.05**	0.07***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
part. per treatment		0.00	0.00		0.00	0.00
		(0.00)	(0.00)		(0.00)	(0.00)
treatments		0.07	0.05		0.06	0.04*
		(0.03)	(0.02)		(0.03)	(0.01)
strength of incentives		1.15	0.91		0.98	0.88
		(1.18)	(0.73)		(0.91)	(0.44)
pages (EE-equivalent)			-0.01			-0.01
			(0.01)			(0.00)
references			0.01			0.01*
			(0.01)			(0.00)
JEL-micro			-0.52			-0.52
			(0.40)			(0.34)
JEL-theory			-0.33			-0.34
			(0.27)			(0.23)
JEL-labor			-0.78**			-0.60
			(0.15)			(0.37)
JEL-econometrics			0.58			0.75*
			(0.23)			(0.21)
JEL-industrial			-1.85***			-1.92**
			(0.08)			(0.25)
JEL-international			n.a.			n.a.
JEL-finance			-1.76**			-1.67**
			(0.38)			(0.39)
JEL-public			-0.41			-0.52*
			(0.35)			(0.13)
JEL-health			-2.22***			-2.26***
			(0.10)			(0.14)
JEL-development			-7.00*			-6.95*
			(1.95)			(2.31)
JEL-other			1.18			1.17
			(0.69)			(0.80)
share fem. auth.			0.24			0.22
			(0.18)			(0.21)
reputation			0.00**			0.00
			(0.00)			(0.00)
number of auth.			-0.11*			-0.15
			(0.04)			(0.09)
Constant	3.78***	2.85***	3.42***	3.68***	2.87***	3.24***
	(0.05)	(0.19)	(0.17)	(0.13)	(0.15)	(0.12)
N	113	112	111	116	115	114
Adjusted R^2	0.34	0.36	0.41	0.36	0.37	0.44

Table 2.7: Impact of social ties on paper quality (full results, only US papers).

Notes. OLS regressions with standard errors clustered by journal in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	dependent variable: asinh(WoS citations)					
	(1)	(2)	(3)	(4)	(5)	(6)
AER	1.72***	1.54***	1.50***	1.52***	1.42***	1.46***
	(0.02)	(0.11)	(0.10)	(0.03)	(0.07)	(0.06)
JEEA	0.92***	0.95***	0.79**	1.07***	1.02***	1.05***
	(0.00)	(0.03)	(0.11)	(0.01)	(0.01)	(0.02)
share US auth.	-0.15***	-0.12**	-0.15**			
	(0.01)	(0.02)	(0.02)			
share US auth. * AER	-0.39***	-0.22*	-0.03			
	(0.00)	(0.07)	(0.16)			
share US auth. * JEEA	0.12***	-0.04	0.22			
	(0.00)	(0.02)	(0.19)			
colleague con.				0.10***	0.04***	-0.10*
				(0.00)	(0.00)	(0.02)
colleague con. * AER				-0.46***	-0.29**	-0.03
				(0.02)	(0.04)	(0.11)
colleague con. * JEEA				-0.56***	-0.54***	-0.94**
				(0.03)	(0.03)	(0.12)
age	0.11***	0.11**	0.11**	0.10***	0.11**	0.11**
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)
part. per treatment		0.00	0.00		0.00	0.00
		(0.00)	(0.00)		(0.00)	(0.00)
treatments		0.04*	0.04		0.04**	0.04
		(0.01)	(0.03)		(0.01)	(0.02)
strength of incentives		0.71***	0.68**		0.76**	0.69**
		(0.07)	(0.10)		(0.09)	(0.14)
pages (EE-equivalent)			-0.02			-0.02
			(0.02)			(0.02)
references			0.01			0.01
			(0.00)			(0.00)
JEL-micro			-0.16			-0.15
			(0.29)			(0.29)
JEL-theory			-0.06			-0.07
			(0.16)			(0.12)
JEL-labor			-0.28			-0.25
			(0.31)			(0.30)
JEL-econometrics			-1.07**			-1.04**
			(0.22)			(0.21)
JEL-industrial			-1.31**			-1.44***
			(0.16)			(0.12)
JEL-international			-3.57**			-3.57**
			(0.45)			(0.37)
JEL-finance			-1.14*			-1.10**
			(0.27)			(0.25)
JEL-public			1.14*			1.08*
			(0.27)			(0.28)
JEL-health			-1.07			-1.12
			(0.73)			(0.63)
JEL-development			-1.34*			-1.18*
			(0.43)			(0.31)
JEL-other			0.30			0.24
			(0.46)			(0.44)
share fem. auth.			-0.05			-0.08
			(0.28)			(0.27)
reputation			0.00			0.00**
			(0.00)			(0.00)
number of auth.			-0.06			-0.05
			(0.04)			(0.04)
Constant	2.19***	1.76**	1.86***	2.19***	1.71***	1.82***
	(0.06)	(0.19)	(0.13)	(0.09)	(0.17)	(0.10)
N	302	300	291	302	300	291
Adjusted R^2	0.43	0.43	0.49	0.42	0.43	0.49

Table 2.8: Impact of social ties on paper quality (full results I, WoS citations).

Notes. OLS regressions with standard errors clustered by journal in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	dependent variable: asinh(WoS citations)					
	(1)	(2)	(3)	(4)	(5)	(6)
AER	0.38***	0.23***	0.73	1.53***	1.44***	1.50***
	(0.01)	(0.01)	(0.31)	(0.04)	(0.07)	(0.05)
JEEA	0.44**	0.31*	-0.04	0.92***	0.88***	0.80***
	(0.07)	(0.07)	(0.27)	(0.00)	(0.02)	(0.04)
minDist	0.01	0.01*	0.02			
	(0.01)	(0.00)	(0.01)			
minDist * AER	0.27***	0.29***	0.18			
	(0.01)	(0.02)	(0.07)			
minDist * JEEA	0.12**	0.15***	0.20*			
	(0.01)	(0.01)	(0.06)			
co-aut. con.				-0.05**	-0.13**	-0.04
				(0.01)	(0.02)	(0.06)
co-aut. con. * AER				-0.87***	-0.94***	-0.85*
				(0.02)	(0.07)	(0.21)
co-aut. con. * JEEA				n.a.	n.a.	n.a.
age	0.10**	0.10**	0.11**	0.10**	0.10**	0.11**
	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
part. per treatment		0.00	0.00		0.00	0.00
		(0.00)	(0.00)		(0.00)	(0.00)
treatments		0.04*	0.04		0.04**	0.04
		(0.01)	(0.02)		(0.01)	(0.02)
strength of incentives		0.88**	0.74*		0.86**	0.76*
		(0.19)	(0.24)		(0.17)	(0.25)
pages (EE-equivalent)			-0.02			-0.02
			(0.02)			(0.02)
references			0.01			0.01
			(0.01)			(0.01)
JEL-micro			-0.15			-0.17
			(0.31)			(0.32)
JEL-theory			-0.05			-0.07
			(0.18)			(0.17)
JEL-labor			-0.53			-0.54
			(0.21)			(0.22)
JEL-econometrics			-1.08**			-1.07**
			(0.25)			(0.23)
JEL-industrial			-1.24**			-1.34**
			(0.19)			(0.15)
JEL-international			-4.33***			-4.38***
			(0.37)			(0.09)
JEL-finance			-1.20**			-1.17**
			(0.14)			(0.20)
JEL-public			1.05**			1.14*
			(0.23)			(0.33)
JEL-health			-1.01			-1.18
			(0.62)			(0.67)
JEL-development			-1.50			-1.40*
			(0.59)			(0.48)
JEL-other			0.23			0.26
			(0.41)			(0.46)
share fem. auth.			-0.09			-0.15
			(0.34)			(0.24)
reputation			0.00**			0.00
			(0.00)			(0.00)
number of auth.			-0.01			-0.03
			(0.01)			(0.03)
Constant	2.17***	1.66**	1.65***	2.20***	1.70***	1.77***
	(0.15)	(0.21)	(0.13)	(0.10)	(0.17)	(0.15)
N	295	293	284	295	293	284
Adjusted R^2	0.43	0.44	0.49	0.42	0.43	0.49

Table 2.9: Impact of social ties on paper quality (full results II, WoS citations).

Notes. OLS regressions with standard errors clustered by journal in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	dependent variable: raw GS citations					
	(1)	(2)	(3)	(4)	(5)	(6)
AER	431.43***	447.15***	452.34***	297.12***	298.45***	328.85***
	(16.27)	(7.73)	(40.97)	(16.78)	(16.92)	(10.47)
JEEA	134.95**	151.80**	157.59**	114.30***	106.25***	112.84**
	(14.63)	(16.16)	(17.90)	(5.05)	(5.09)	(12.05)
share US auth.	-26.83	-25.55	-35.04			
	(22.53)	(24.79)	(39.29)			
share US auth. * AER	-265.66***	-275.78***	-212.96			
	(1.49)	(13.51)	(90.62)			
share US auth. * JEEA	-98.51***	-134.18***	-104.53			
	(2.16)	(4.12)	(50.83)			
colleague con.				3.56	-1.05	-10.02
				(1.40)	(0.73)	(5.89)
colleague con. * AER				-181.90***	-179.31***	-124.36
				(16.85)	(15.81)	(58.75)
colleague con. * JEEA				-69.23	-47.13	-57.08
				(29.94)	(34.38)	(77.48)
age	13.41	12.95	10.60	11.90	12.05	10.34*
	(10.65)	(10.14)	(3.95)	(9.03)	(8.79)	(2.91)
part. per treatment		-0.33	-0.20		-0.13	0.01
		(0.23)	(0.12)		(0.15)	(0.09)
treatments		3.95	0.98		6.56**	3.61
		(4.32)	(10.12)		(1.17)	(7.20)
strength of incentives		10.23	20.88		49.06	47.62
		(16.96)	(50.36)		(60.57)	(84.42)
pages (EE-equivalent)			0.01			-0.30
			(3.42)			(3.30)
references			-0.23			0.09
			(2.47)			(2.29)
JEL-micro			30.86			22.79
			(12.11)			(12.90)
JEL-theory			-92.31			-130.72
			(76.75)			(100.24)
JEL-labor			-91.23			-64.38
			(75.11)			(52.85)
JEL-econometrics			14.59			7.99
			(42.27)			(42.76)
JEL-industrial			-68.60*			-77.44**
			(19.54)			(10.52)
JEL-international			-16.77			28.38
			(64.63)			(91.64)
JEL-finance			-99.28			-92.10
			(73.51)			(72.12)
JEL-public			735.21			771.16
			(741.66)			(786.38)
JEL-health			-309.64			-326.86
			(206.72)			(199.39)
JEL-development			-57.48			-62.05**
			(69.63)			(11.14)
JEL-other			157.19			136.95
			(179.89)			(182.67)
share fem. auth.			33.79			20.85
			(47.41)			(35.27)
reputation			0.01			0.03
			(0.03)			(0.05)
number of auth.			-23.37			-18.36
			(36.54)			(32.45)
Constant	-62.03	-58.84	-8.91	-56.84	-92.20	-58.42
	(98.54)	(73.48)	(63.51)	(90.57)	(102.05)	(32.99)
N	329	327	318	329	327	318
Adjusted R^2	0.19	0.19	0.29	0.16	0.16	0.27

Table 2.10: Impact of social ties on paper quality (full results I, raw citations).

Notes. OLS regressions with standard errors clustered by journal in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	dependent variable: raw GS citations					
	(1)	(2)	(3)	(4)	(5)	(6)
AER	-270.64***	-277.34***	-178.93**	288.71***	288.58***	321.48***
	(0.20)	(14.68)	(26.32)	(16.07)	(18.20)	(9.32)
JEEA	21.05	31.47	65.05	95.95**	93.71***	101.27*
	(28.85)	(37.53)	(112.13)	(10.22)	(8.97)	(23.56)
minDist	-1.02***	-0.12	-7.10			
	(0.02)	(0.72)	(7.07)			
minDist * AER	133.46***	135.24***	116.06***			
	(3.19)	(0.95)	(7.29)			
minDist * JEEA	18.41*	16.02	5.63			
	(4.82)	(6.71)	(24.75)			
co-aut. con.				-16.93	-30.90	39.05
				(12.39)	(18.98)	(44.79)
co-aut. con. * AER				-223.25**	-233.66***	-266.74***
				(22.84)	(9.67)	(21.54)
co-aut. con. * JEEA				n.a.	n.a.	n.a.
age	10.61	10.47	9.06**	11.34	11.42	9.61**
	(7.20)	(6.70)	(1.41)	(8.12)	(7.79)	(1.89)
part. per treatment		-0.17	-0.04		-0.13	0.00
		(0.18)	(0.12)		(0.14)	(0.11)
treatments		5.45*	2.62		6.92**	2.86
		(1.76)	(7.95)		(1.55)	(7.83)
strength of incentives		101.21	94.44		79.75	81.38
		(112.28)	(126.63)		(91.69)	(114.76)
pages (EE-equivalent)			0.79			0.38
			(4.35)			(3.83)
references			-0.75			-0.04
			(3.24)			(2.44)
JEL-micro			26.65			16.72
			(9.80)			(15.75)
JEL-theory			-105.00			-118.40
			(79.17)			(89.20)
JEL-labor			-52.51*			-74.98
			(16.91)			(28.71)
JEL-econometrics			25.03			11.01
			(54.86)			(44.34)
JEL-industrial			-57.23**			-64.83**
			(9.85)			(12.06)
JEL-international			158.10			43.78
			(241.35)			(125.47)
JEL-finance			-105.96			-88.83
			(91.85)			(65.15)
JEL-public			775.92			802.50
			(784.87)			(803.92)
JEL-health			-303.28			-357.24
			(172.04)			(199.79)
JEL-development			-109.35			-91.50**
			(48.25)			(18.40)
JEL-other			137.83			154.11
			(178.97)			(190.73)
share fem. auth.			38.55			15.20
			(53.91)			(32.48)
reputation			0.02			0.02
			(0.04)			(0.04)
number of auth.			-14.17			-18.11
			(25.04)			(31.08)
Constant	-38.80	-86.09	-33.26	-50.52	-95.76	-70.31
	(72.40)	(97.76)	(50.40)	(80.96)	(103.87)	(29.70)
N	321	319	310	321	319	310
Adjusted R^2	0.20	0.20	0.29	0.15	0.15	0.27

Table 2.11: Impact of social ties on paper quality (full results II, raw citations).

Notes. OLS regressions with standard errors clustered by journal in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	dependent variable: raw GS citations					
	(1)	(2)	(3)	(4)	(5)	(6)
AER	11.02**	-50.43	-62.76	205.71***	188.12***	177.86
	(1.29)	(57.72)	(110.95)	(4.01)	(15.39)	(75.12)
JEEA	-246.62***	-129.66	-190.77	41.88***	6.46	30.98*
	(10.71)	(167.88)	(104.68)	(3.61)	(32.79)	(9.32)
minDist	-1.53**	0.49	-5.33			
	(0.36)	(3.58)	(3.24)			
minDist * AER	43.08***	55.87*	58.35*			
	(0.04)	(13.00)	(14.88)			
minDist * JEEA	63.10***	25.22	46.41			
	(2.60)	(51.03)	(18.82)			
any-connection				22.71	-3.12	2.30
				(9.73)	(13.66)	(14.24)
any-connection * AER				-98.75**	-67.08**	-40.87
				(12.13)	(12.93)	(39.00)
any-connection * JEEA				n.a.	n.a.	n.a.
age	2.94*	3.72**	5.36	4.91	5.98	8.09
	(0.69)	(0.63)	(1.94)	(3.34)	(2.98)	(6.12)
part. per treatment		0.27	0.44		0.40*	0.51
		(0.26)	(0.16)		(0.11)	(0.24)
treatments		15.30	16.27		14.92	14.00
		(11.83)	(16.36)		(10.12)	(12.33)
strength of incentives		230.44	239.56		174.77	194.69
		(291.40)	(297.02)		(203.53)	(180.59)
pages (EE-equivalent)			0.77			0.59
			(2.98)			(2.51)
references			0.43			1.57**
			(0.99)			(0.24)
JEL-micro			-17.35			-1.61
			(21.58)			(42.80)
JEL-theory			-20.22			-40.35*
			(17.18)			(11.58)
JEL-labor			-70.49			-84.30
			(75.42)			(120.14)
JEL-econometrics			17.90			32.08
			(28.57)			(42.41)
JEL-industrial			-116.63			-148.48
			(55.47)			(92.63)
JEL-international			n.a.			n.a.
JEL-finance			-157.40			-204.02
			(265.27)			(260.58)
JEL-public			-34.85			-65.30
			(18.16)			(48.98)
JEL-health			-383.77			-395.10
			(158.62)			(175.15)
JEL-development			-449.65**			-355.63**
			(65.82)			(63.15)
JEL-other			197.88			163.93
			(101.00)			(86.28)
share fem. auth.			108.40			80.51
			(92.09)			(73.85)
reputation			-0.06			-0.06
			(0.09)			(0.09)
number of auth.			-11.85			-22.63
			(18.72)			(39.43)
Constant	27.21**	-129.08	-129.46	-3.25	-138.73	-158.29
	(6.07)	(128.03)	(158.09)	(37.55)	(124.25)	(140.65)
N	113	112	111	116	115	114
Adjusted R^2	0.20	0.23	0.25	0.22	0.24	0.26

Table 2.12: Impact of social ties on paper quality (full results, only US papers, raw citations).

Notes. OLS regressions with standard errors clustered by journal in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

2.5.2 List with all papers included in the data set

American Economic Review

1. Friedman, D. (1998), Monty Hall's three doors: Construction and deconstruction of a choice anomaly. *American Economic Review*, *88*(4), 933-946.
2. Ho, T. H., Camerer, C., & Weigelt, K. (1998). Iterated dominance and iterated best response in experimental" p-beauty contests". *American Economic Review*, *88*(4), 947-969.
3. Durham, Y., Hirshleifer, J., & Smith, V. L. (1998). Do the rich get richer and the poor poorer? Experimental tests of a model of power. *American Economic Review*, *88*(4), 970-983.
4. Rich, C. S., & Friedman, D. (1998). The matching market institution: A laboratory investigation. *American Economic Review*, *88*(5), 1311-1322.
5. Blume, A., DeJong, D. V., Kim, Y. G., & Sprinkle, G. B. (1998). Experimental evidence on the evolution of meaning of messages in sender-receiver games. *American Economic Review*, *88*(5), 1323-1340.
6. Corns, A., & Schotter, A. (1999). Can affirmative action be cost effective? An experimental examination of price-preference auctions. *American Economic Review*, *89*(1), 291-305.
7. Camerer, C., & Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review*, *89*(1), 306-318.
8. Cummings, R. G., & Taylor, L. O. (1999). Unbiased value estimates for environmental goods: a cheap talk design for the contingent valuation method. *American Economic Review*, *89*(3), 649-665.
9. Capra, C. M., Goeree, J. K., Gomez, R., & Holt, C. A. (1999). Anomalous behavior in a traveler's dilemma?. *American Economic Review*, *89*(3), 678-690.
10. Duffy, J., & Ochs, J. (1999). Emergence of money as a medium of exchange: An experimental study. *American Economic Review*, *89*(4), 847-877.
11. Falkinger, J., Fehr, E., Gächter, S., & Winter-Ember, R. (2000). A simple mechanism for the efficient provision of public goods: Experimental evidence. *American Economic Review*, *90*(1), 247-264.
12. Fehr, E., & Gachter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, *90*(4), 980-994.
13. James, D., & Isaac, R. M. (2000). Asset markets: How they are affected by tournament incentives for individuals. *American Economic Review*, *90*(4), 995-1004.
14. Konow, J. (2000). Fair shares: Accountability and cognitive dissonance in allocation decisions. *American economic review*, *90*(4), 1072-1091.
15. Fehr, E., & Tyran, J. R. (2001). Does money illusion matter?. *American Economic Review*, *91*(5), 1239-1262.
16. Goeree, J. K., & Holt, C. A. (2001). Ten little treasures of game theory and ten intuitive contradictions. *American Economic Review*, *91*(5), 1402-1422.
17. Deck, C. A. (2001). A test of game-theoretic and behavioral models of play in exchange and insurance environments. *American Economic Review*, *91*(5), 1546-1555.
18. Goeree, K., & Offerman, T. (2002). Efficiency in auctions with private and common values: An experimental study. *American Economic Review*, *92*(3), 625-643.
19. Houser, D., & Kurzban, R. (2002). Revisiting kindness and confusion in public goods experiments. *American Economic Review*, *92*(4), 1062-1069.
20. Cherry, T. L., Frykblom, P., & Shogren, J. F. (2002). Hardnose the dictator. *American Economic Review*, *92*(4), 1218-1221.

21. Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American economic review*, *92*(5), 1644-1655.
22. Chen, Y., & Snmez, T. (2002). Improving efficiency of on-campus housing: An experimental study. *American Economic Review*, *92*(5), 1669-1686.
23. Masclet, D., Noussair, C., Tucker, S., & Villeval, M. C. (2003). Monetary and nonmonetary punishment in the voluntary contributions mechanism. *American Economic Review*, *93*(1), 366-380.
24. Andreoni, J., Castillo, M., & Petrie, R. (2003). What do bargainers' preferences look like? Experiments with a convex ultimatum game. *American Economic Review*, *93*(3), 672-685.
25. Andreoni, J., Harbaugh, W., & Vesterlund, L. (2003). The carrot or the stick: Rewards, punishments, and cooperation. *American Economic Review*, *93*(3), 893-902.
26. Phillips, O. R., Menkhaus, D. J., & Coatney, K. T. (2003). Collusive practices in repeated English auctions: Experimental evidence on bidding rings. *American Economic Review*, *93*(3), 965-979.
27. Çelen, B., & Kariv, S. (2004). Distinguishing informational cascades from herd behavior in the laboratory. *American Economic Review*, *94*(3), 484-498.
28. List, J. A., Berrens, R. P., Bohara, A. K., & Kerkvliet, J. (2004). Examining the role of social isolation on stated preferences. *American Economic Review*, *94*(3), 741-752.
29. Engelmann, D., & Strobel, M. (2004). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *American economic review*, *94*(4), 857-869.
30. Chen, Y., & Gazzale, R. (2004). When does learning in games generate convergence to Nash equilibria? The role of supermodularity in an experimental setting. *American Economic Review*, *94*(5), 1505-1535.
31. Gneezy, U. (2005). Deception: The role of consequences. *American Economic Review*, *95*(1), 384-394.
32. Cooper, D. J., & Kagel, J. H. (2005). Are two heads better than one? Team versus individual play in signaling games. *American Economic Review*, *95*(3), 477-509.
33. Duffy, J., & Fisher, E. O. N. (2005). Sunspots in the Laboratory. *American Economic Review*, *95*(3), 510-529.
34. Plott, C. R., & Zeiler, K. (2005). The willingness to pay-willingness to accept gap, the" endowment effect," subject misconceptions, and experimental procedures for eliciting valuations. *American Economic Review*, *95*(3), 530-545.
35. Holt, C. A., & Laury, S. K. (2005). Risk aversion and incentive effects: New data without order effects. *American Economic Review*, *95*(3), 902-912.
36. Charness, G., & Levin, D. (2005). When optimal choices feel wrong: A laboratory study of Bayesian updating, complexity, and affect. *American Economic Review*, *95*(4), 1300-1309.
37. Cipriani, M., & Guarino, A. (2005). Herd behavior in a laboratory financial market. *American Economic Review*, *95*(5), 1427-1443.
38. Bó, P. D. (2005). Cooperation under the shadow of the future: experimental evidence from infinitely repeated games. *American economic review*, *95*(5), 1591-1604.
39. Dufwenberg, M., Lindqvist, T., & Moore, E. (2005). Bubbles and experience: An experiment. *American Economic Review*, *95*(5), 1731-1737.
40. Weber, R. A. (2006). Managing growth to achieve efficient coordination in large groups. *American Economic Review*, *96*(1), 114-126.

41. Mobius, M. M., & Rosenblat, T. S. (2006). Why beauty matters. *American Economic Review*, *96*(1), 222-235.
42. Brandts, J., & Cooper, D. J. (2006). A change would do you good.... An experimental study on how to overcome coordination failure in organizations. *American Economic Review*, *96*(3), 669-693.
43. Bereby-Meyer, Y., & Roth, A. E. (2006). The speed of learning in noisy games: Partial reinforcement and the sustainability of cooperation. *American Economic Review*, *96*(4), 1029-1042.
44. Gabaix, X., Laibson, D., Moloche, G., & Weinberg, S. (2006). Costly information acquisition: Experimental analysis of a boundedly rational model. *American Economic Review*, *96*(4), 1043-1068.
45. Falk, A., & Kosfeld, M. (2006). The hidden costs of control. *American Economic Review*, *96*(5), 1611-1630.
46. Butler, D. J., & Loomes, G. C. (2007). Imprecision as an account of the preference reversal phenomenon. *American Economic Review*, *97*(1), 277-297.
47. Cappelen, A. W., Hole, A. D., Sørensen, E. Ø., & Tungodden, B. (2007). The pluralism of fairness ideals: An experimental approach. *American Economic Review*, *97*(3), 818-827.
48. Casari, M., Ham, J. C., & Kagel, J. H. (2007). Selection bias, demographic effects, and ability effects in common value auction experiments. *American Economic Review*, *97*(4), 1278-1304.
49. Filiz-Ozbay, E., & Ozbay, E. Y. (2007). Auctions with anticipated regret: Theory and experiment. *American Economic Review*, *97*(4), 1407-1418.
50. Healy, P. J. (2007). Group reputations, stereotypes, and cooperation in a repeated labor market. *American Economic Review*, *97*(5), 1751-1773.
51. Fisman, R., Kariv, S., & Markovits, D. (2007). Individual preferences for giving. *American Economic Review*, *97*(5), 1858-1876.
52. Choi, S., Fisman, R., Gale, D., & Kariv, S. (2007). Consistency and heterogeneity of individual behavior under uncertainty. *American Economic Review*, *97*(5), 1921-1938.
53. Charness, G., Rigotti, L., & Rustichini, A. (2007). Individual behavior and group membership. *American Economic Review*, *97*(4), 1340-1352.
54. Haruvy, E., Lahav, Y., & Noussair, C. N. (2007). Traders' expectations in asset markets: experimental evidence. *American Economic Review*, *97*(5), 1901-1920.
55. Post, T., Van den Assem, M. J., Baltussen, G., & Thaler, R. H. (2008). Deal or no deal? decision making under risk in a large-payoff game show. *American Economic Review*, *98*(1), 38-71.
56. Hussam, R. N., Porter, D., & Smith, V. L. (2008). Thar she blows: Can bubbles be rekindled with experienced subjects?. *American Economic Review*, *98*(3), 924-37.
57. Selten, R., & Chmura, T. (2008). Stationary concepts for experimental 2x2-games. *American Economic Review*, *98*(3), 938-66.
58. Kimbrough, E. O., Smith, V. L., & Wilson, B. J. (2008). Historical property rights, sociality, and the emergence of impersonal exchange in long-distance trade. *American Economic Review*, *98*(3), 1009-39.
59. Crawford, V. P., Gneezy, U., & Rottenstreich, Y. (2008). The power of focal points is limited: Even minute payoff asymmetry may yield large coordination failures. *American Economic Review*, *98*(4), 1443-58.
60. Hargreaves Heap, S. P., & Zizzo, D. J. (2009). The value of groups. *American Economic Review*, *99*(1), 295-323.

61. Chen, Y., & Li, S. X. (2009). Group identity and social preferences. *American Economic Review*, *99*(1), 431-57.
62. Camera, G., & Casari, M. (2009). Cooperation among strangers under the shadow of the future. *American Economic Review*, *99*(3), 979-1005.
63. Blume, A., Duffy, J., & Franco, A. M. (2009). Decentralized organizational learning: an experimental investigation. *American Economic Review*, *99*(4), 1178-1205.
64. Kosfeld, M., Okada, A., & Riedl, A. (2009). Institution formation in public goods games. *American Economic Review*, *99*(4), 1335-55.
65. Ivanov, A., Levin, D., & Peck, J. (2009). Hindsight, foresight, and insight: an experimental study of a small-market investment game with common and private values. *American Economic Review*, *99*(4), 1484-1507.
66. Rabin, M., & Weizsäcker, G. (2009). Narrow bracketing and dominated choices. *American Economic Review*, *99*(4), 1508-43.
67. Landeo, C. M., & Spier, K. E. (2009). Naked exclusion: an experimental study of contracts with externalities. *American Economic Review*, *99*(5), 1850-77.
68. Ho, T. H., & Su, X. (2009). Peer-induced fairness in games. *American Economic Review*, *99*(5), 2022-49.
69. Daniel, T. E., Gisches, E. J., & Rapoport, A. (2009). Departure times in Y-shaped traffic networks with multiple bottlenecks. *American Economic Review*, *99*(5), 2149-76.
70. Abbink, K., Brandts, J., Herrmann, B., & Orzen, H. (2010). Intergroup conflict and intra-group punishment in an experimental contest game. *American Economic Review*, *100*(1), 420-47.
71. Carpenter, J., Matthews, P. H., & Schirm, J. (2010). Tournaments and office politics: Evidence from a real effort experiment. *American Economic Review*, *100*(1), 504-17.
72. Fischbacher, U., & Gachter, S. (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American economic review*, *100*(1), 541-56.
73. Bushong, B., King, L. M., Camerer, C. F., & Rangel, A. (2010). Pavlovian processes in consumer choice: The physical presence of a good increases willingness-to-pay. *American Economic Review*, *100*(4), 1556-71.
74. Anderson, S. T., Friedman, D., & Oprea, R. (2010). Preemption games: Theory and experiment. *American Economic Review*, *100*(4), 1778-1803.
75. Hamman, J. R., Loewenstein, G., & Weber, R. A. (2010). Self-interest through delegation: An additional rationale for the principal-agent relationship. *American Economic Review*, *100*(4), 1826-46.
76. Calsamiglia, C., Haeringer, G., & Klijn, F. (2010). Constrained school choice: An experimental study. *American Economic Review*, *100*(4), 1860-74.
77. Feri, F., Irlenbusch, B., & Sutter, M. (2010). Efficiency gains from team-based coordination—large-scale experimental evidence. *American Economic Review*, *100*(4), 1892-1912.
78. Benjamin, D. J., Choi, J. J., & Strickland, A. J. (2010). Social identity and preferences. *American Economic Review*, *100*(4), 1913-28.
79. Grosskopf, B., & Sarin, R. (2010). Is reputation good or bad? An experiment. *American Economic Review*, *100*(5), 2187-2204.
80. Dal Bó, P., Foster, A., & Putterman, L. (2010). Institutions and behavior: Experimental evidence on the effects of democracy. *American Economic Review*, *100*(5), 2205-29.

81. Cabrales, A., Miniaci, R., Piovesan, M., & Ponti, G. (2010). Social preferences and strategic uncertainty: an experiment on markets and contracts. *American Economic Review*, *100*(5), 2261-78.
82. Ellman, M., & Pezanis-Christou, P. (2010). Organizational structure, communication, and group ethics. *American Economic Review*, *100*(5), 2478-91.
83. Dal Bó, P., & Fréchette, G. R. (2011). The evolution of cooperation in infinitely repeated games: Experimental evidence. *American Economic Review*, *101*(1), 411-29.
84. Abeler, J., Falk, A., Goette, L., & Huffman, D. (2011). Reference points and effort provision. *American Economic Review*, *101*(2), 470-92.
85. Fehr, E., Hart, O., & Zehnder, C. (2011). Contracts as reference points—experimental evidence. *American Economic Review*, *101*(2), 493-525.
86. Dulleck, U., Kerschbamer, R., & Sutter, M. (2011). The economics of credence goods: An experiment on the role of liability, verifiability, reputation, and competition. *American Economic Review*, *101*(2), 526-55.
87. Dohmen, T., & Falk, A. (2011). Performance pay and multidimensional sorting: Productivity, preferences, and gender. *American economic review*, *101*(2), 556-90.
88. Huck, S., Seltzer, A. J., & Wallace, B. (2011). Deferred compensation in multiperiod labor contracts: an experimental test of Lazear's model. *American Economic Review*, *101*(2), 819-43.
89. Kogan, S., Kwasnica, A. M., & Weber, R. A. (2011). Coordination in the presence of asset markets. *American Economic Review*, *101*(2), 927-47.
90. Brown, M., Flinn, C. J., & Schotter, A. (2011). Real-time search in the laboratory and the market. *American Economic Review*, *101*(2), 948-74.
91. Charness, G., & Dufwenberg, M. (2011). Participation. *American Economic Review*, *101*(4), 1211-37.
92. Eckel, C. C., & Petrie, R. (2011). Face value. *American Economic Review*, *101*(4), 1497-1513.
93. Chen, R., & Chen, Y. (2011). The potential of social identity for equilibrium selection. *American Economic Review*, *101*(6), 2562-89.
94. Caplin, A., Dean, M., & Martin, D. (2011). Search and satisficing. *American Economic Review*, *101*(7), 2899-2922.
95. Ifcher, J., & Zarghamee, H. (2011). Happiness and time preference: The effect of positive affect in a random-assignment experiment. *American Economic Review*, *101*(7), 3109-29.
96. Crockett, S., Oprea, R., & Plott, C. (2011). Extreme walrasian dynamics: The gale example in the lab. *American Economic Review*, *101*(7), 3196-3220.
97. Friedman, D., & Oprea, R. (2012). A continuous dilemma. *American Economic Review*, *102*(1), 337-63.
98. Gill, D., & Prowse, V. (2012). A structural analysis of disappointment aversion in a real effort competition. *American Economic Review*, *102*(1), 469-503.
99. Fudenberg, D., Rand, D. G., & Dreber, A. (2012). Slow to anger and fast to forgive: Cooperation in an uncertain world. *American Economic Review*, *102*(2), 720-49.
100. Bartling, B., Fehr, E., & Schmidt, K. M. (2012). Screening, competition, and job design: Economic origins of good jobs. *American Economic Review*, *102*(2), 834-64.
101. Kirchler, M., Huber, J., & Stöckl, T. (2012). Thar she bursts: Reducing confusion reduces bubbles. *American Economic Review*, *102*(2), 865-83.
102. Kessler, J. B., & Roth, A. E. (2012). Organ allocation policy and the decision to donate. *American Economic Review*, *102*(5), 2018-47.

103. Charness, G., Cobo-Reyes, R., Jiménez, N., Lacomba, J. A., & Lagos, F. (2012). The hidden advantage of delegation: Pareto improvements in a gift exchange game. *American Economic Review*, *102*(5), 2358-79.
104. Ambrus, A., & Greiner, B. (2012). Imperfect public monitoring with costly punishment: An experimental study. *American Economic Review*, *102*(7), 3317-32.
105. Andreoni, J., & Sprenger, C. (2012). Estimating time preferences from convex budgets. *American Economic Review*, *102*(7), 3333-56.
106. Andreoni, J., & Sprenger, C. (2012). Risk preferences are not time preferences. *American Economic Review*, *102*(7), 3357-76.
107. Arad, A., & Rubinstein, A. (2012). The 11-20 money request game: A level-k reasoning study. *American Economic Review*, *102*(7), 3561-73.
108. Brock, J. M., Lange, A., & Ozbay, E. Y. (2013). Dictating the risk: Experimental evidence on giving in risky environments. *American Economic Review*, *103*(1), 415-37.
109. Gibson, R., Tanner, C., & Wagner, A. F. (2013). Preferences for truthfulness: Heterogeneity among and within individuals. *American Economic Review*, *103*(1), 532-48.
110. Fehr, E., Herz, H., & Wilkening, T. (2013). The lure of authority: Motivation and incentive effects of power. *American Economic Review*, *103*(4), 1325-59.
111. Maniadis, Z., Tufano, F., & List, J. A. (2014). One swallow doesn't make a summer: New evidence on anchoring effects. *American Economic Review*, *104*(1), 277-90.
112. Duffy, J., & Puzzello, D. (2014). Gift exchange versus monetary exchange: Theory and evidence. *American economic review*, *104*(6), 1735-76.
113. Isoni, A., Poulsen, A., Sugden, R., & Tsutsui, K. (2014). Efficiency, equality, and labeling: An experimental investigation of focal points in explicit bargaining. *American Economic Review*, *104*(10), 3256-87.
114. De Clippel, G., Eliaz, K., & Knight, B. (2014). On the selection of arbitrators. *American Economic Review*, *104*(11), 3434-58.
115. Jackson, M. O., & Yariv, L. (2014). Present bias and collective dynamic choice in the lab. *American Economic Review*, *104*(12), 4184-4204.
116. Eckel, C. C., & Füllbrunn, S. C. (2015). Thar she blows? Gender, competition, and bubbles in experimental asset markets. *American Economic Review*, *105*(2), 906-20.
117. Di Tella, R., Perez-Truglia, R., Babino, A., & Sigman, M. (2015). Conveniently upset: Avoiding altruism by distorting beliefs about others' altruism. *American Economic Review*, *105*(11), 3416-42.
118. Attema, A. E., Bleichrodt, H., Gao, Y., Huang, Z., & Wakker, P. P. (2016). Measuring discounting without measuring utility. *American Economic Review*, *106*(6), 1476-94.
119. Malmendier, U., & Schmidt, K. M. (2017). You owe me. *American Economic Review*, *107*(2), 493-526.
120. Ottoni-Wilhelm, M., Vesterlund, L., & Xie, H. (2017). Why do people give? Testing pure and impure altruism. *American Economic Review*, *107*(11), 3617-33.
121. Gneezy, U., Kajackaite, A., & Sobel, J. (2018). Lying Aversion and the Size of the Lie. *American Economic Review*, *108*(2), 419-53.

Experimental Economics

122. Hey, J. D., & Di Cagno, D. (1998). Sequential markets: An experimental investigation of Clower's dual-decision hypothesis. *Experimental Economics*, *1*(1), 63-85.

123. Bornstein, G., & Yaniv, I. (1998). Individual and group behavior in the ultimatum game: are groups more "rational" players?. *Experimental Economics*, *1*(1), 101-108.
124. Cubitt, R. P., Starmer, C., & Sugden, R. (1998). On the validity of the random lottery incentive system. *Experimental Economics*, *1*(2), 115-131.
125. Wilson, B. J. (1998). What collusion? Unilateral market power as a catalyst for countercyclical markups. *Experimental Economics*, *1*(2), 133-145.
126. Isaac, R. M., & Walker, J. M. (1998). Nash as an organizing principle in the voluntary provision of public goods: Experimental evidence. *Experimental economics*, *1*(3), 191-206.
127. Bolton, G. E., Brandts, J., & Ockenfels, A. (1998). Measuring motivations for the reciprocal responses observed in a simple dilemma game. *Experimental Economics*, *1*(3), 207-219.
128. Rapoport, A., Daniel, T. E., & Seale, D. A. (1998). Reinforcement-based adaptive learning in asymmetric two-person bargaining with incomplete information. *Experimental Economics*, *1*(3), 221-253.
129. Chan, K. S., Mestelman, S., Moir, R., & Muller, R. A. (1999). Heterogeneity and the voluntary provision of public goods. *Experimental Economics*, *2*(1), 5-30.
130. Herne, K. (1999). The effects of decoy gambles on individual choice. *Experimental Economics*, *2*(1), 31-40.
131. Berninghaus, S. K., Ehrhart, K. M., & Keser, C. (1999). Continuous-time strategy selection in linear population games. *Experimental Economics*, *2*(1), 41-57.
132. Sherstyuk, K. (1999). Collusion without conspiracy: An experimental study of one-sided auctions. *Experimental Economics*, *2*(1), 59-75.
133. Cason, T. N., & Friedman, D. (1999). Learning in a laboratory market with random supply and demand. *Experimental Economics*, *2*(1), 77-98.
134. Coller, M., & Williams, M. B. (1999). Eliciting individual discount rates. *Experimental Economics*, *2*(2), 107-127.
135. Van Huyck, J., Rankin, F., & Battalio, R. (1999). What Does it Take to Eliminate the use of a Strategy Strictly Dominated by a Mixture?. *Experimental economics*, *2*(2), 129-150.
136. Cox, J. C., & Oaxaca, R. L. (2000). Good news and bad news: Search from unknown wage offer distributions. *Experimental Economics*, *2*(3), 197-225.
137. Brandts, J., & Charness, G. (2000). Hot vs. cold: Sequential responses and preference stability in experimental games. *Experimental Economics*, *2*(3), 227-238.
138. Croson, R. T., & Marks, M. B. (2000). Step returns in threshold public goods: A meta-and experimental analysis. *Experimental Economics*, *2*(3), 239-259.
139. Hoffman, E., McCabe, K., & Smith, V. (2000). The impact of exchange context on the activation of equity in ultimatum games. *Experimental Economics*, *3*(1), 5-9.
140. List, J. A., & Cherry, T. L. (2000). Learning to accept in ultimatum games: Evidence from an experimental design that generates low offers. *Experimental Economics*, *3*(1), 11-29.
141. Isaac, R. M., & James, D. (2000). Robustness of the incentive compatible combinatorial auction. *Experimental Economics*, *3*(1), 31-53.
142. Cookson, R. (2000). Framing effects in public goods experiments. *Experimental Economics*, *3*(1), 55-79.
143. Ortmann, A., Fitzgerald, J., & Boeing, C. (2000). Trust, reciprocity, and social history: A re-examination. *Experimental Economics*, *3*(1), 81-100.
144. Eckel, C. C., & Grossman, P. J. (2000). Volunteers and pseudo-volunteers: The effect of recruitment method in dictator experiments. *Experimental Economics*, *3*(2), 107-120.

145. Allsopp, L., & Hey, J. D. (2000). Two experiments to test a model of herd behaviour. *Experimental Economics*, *3*(2), 121-136.
146. Anderhub, V., Güth, W., Müller, W., & Strobel, M. (2000). An experimental analysis of intertemporal allocation behavior. *Experimental Economics*, *3*(2), 137-152.
147. Seale, D. A., & Rapoport, A. (2000). Elicitation of strategy profiles in large group coordination games. *Experimental Economics*, *3*(2), 153-179.
148. McKelvey, R. D., & Page, T. (2000). An experimental study of the effect of private information in the Coase theorem. *Experimental Economics*, *3*(3), 187-213.
149. Bardsley, N. (2000). Control without deception: Individual behaviour in free-riding experiments revisited. *Experimental Economics*, *3*(3), 215-240.
150. Engelmann, D., & Strobel, M. (2000). The false consensus effect disappears if representative information and monetary incentives are given. *Experimental Economics*, *3*(3), 241-260.
151. Harstad, R. M. (2000). Dominant strategy adoption and bidders' experience with pricing rules. *Experimental economics*, *3*(3), 261-280.
152. Hey, J. D. (2001). Does repetition improve consistency?. *Experimental Economics*, *4*(1), 5-54.
153. Krahnen, J. P., & Weber, M. (2001). Marketmaking in the laboratory: Does competition matter?. *Experimental Economics*, *4*(1), 55-85.
154. Noussair, C., Robin, S., & Ruffieux, B. (2001). Price bubbles in laboratory asset markets with constant fundamental values. *Experimental Economics*, *4*(1), 87-105.
155. Dickinson, D. L. (2001). The carrot vs. the stick in work team motivation. *Experimental Economics*, *4*(1), 107-124.
156. Willinger, M., & Ziegelmeyer, A. (2001). Strength of the social dilemma in a public goods experiment: an exploration of the error hypothesis. *Experimental Economics*, *4*(2), 131-144.
157. McDaniel, T. M., & Rutström, E. E. (2001). Decision making costs and problem solving performance. *Experimental Economics*, *4*(2), 145-161.
158. Cox, J. C., Dinkin, S., & Swarthout, J. T. (2001). Endogenous entry and exit in common value auctions. *Experimental Economics*, *4*(2), 163-181.
159. Buckley, N., Chan, K. S., Chowhan, J., Mestelman, S., & Shehata, M. (2001). Value orientations, income and displacement effects, and voluntary contributions. *Experimental Economics*, *4*(2), 183-195.
160. Kagel, J. H., & Wolfe, K. W. (2001). Tests of fairness models based on equity considerations in a three-person ultimatum game. *Experimental Economics*, *4*(3), 203-219.
161. Weber, R. A. (2001). Behavior and learning in the “dirty faces” game. *Experimental Economics*, *4*(3), 229-242.
162. Phillips, O. R., Menkhaus, D. J., & Krogmeier, J. L. (2001). Laboratory behavior in spot and forward auction markets. *Experimental Economics*, *4*(3), 243-256.
163. Knetsch, J. L., Tang, F. F., & Thaler, R. H. (2001). The endowment effect and repeated market trials: Is the Vickrey auction demand revealing?. *Experimental economics*, *4*(3), 257-269.
164. Anderhub, V., Gächter, S., & Königstein, M. (2002). Efficient contracting and fair play in a simple principal-agent experiment. *Experimental Economics*, *5*(1), 5-27.
165. Bornstein, G., & Gneezy, U. (2002). Price competition between teams. *Experimental Economics*, *5*(1), 29-38.
166. Cooper, D. J., & Stockman, C. K. (2002). Learning to punish: experimental evidence from a sequential step-level public goods game. *Experimental Economics*, *5*(1), 39-51.

167. Cason, T. N., Saijo, T., & Yamato, T. (2002). Voluntary participation and spite in public good provision experiments: an international comparison. *Experimental Economics*, *5*(2), 133-153.
168. Archibald, G., & Wilcox, N. T. (2002). A New Variant of the Winner's Curse in a Coasian Contracting Game. *Experimental Economics*, *5*(2), 155-172.
169. Brewer, P. J., Huang, M., Nelson, B., & Plott, C. R. (2002). On the behavioral foundations of the law of supply and demand: Human convergence and robot randomness. *Experimental economics*, *5*(3), 179-208.
170. Swope, K. J. (2002). An experimental investigation of excludable public goods. *Experimental Economics*, *5*(3), 209-222.
171. Clark, J. (2002). House money effects in public good experiments. *Experimental Economics*, *5*(3), 223-231.
172. Roelofs, M. R. (2002). Common value auctions with default: an experimental approach. *Experimental Economics*, *5*(3), 233-252.
173. Zwick, R., & Rapoport, A. (2002). Tacit coordination in a decentralized market entry game with fixed capacity. *Experimental Economics*, *5*(3), 253-272.
174. Muller, R. A., & Sadanand, A. (2003). Order of play, forward induction, and presentation effects in two-person games. *Experimental Economics*, *6*(1), 5-25.
175. Anderhub, V., Güth, W., Kamecke, U., & Normann, H. T. (2003). Capacity choices and price competition in experimental markets. *Experimental Economics*, *6*(1), 27-52.
176. Devetag, G. (2003). Coordination and information in critical mass games: an experimental study. *Experimental economics*, *6*(1), 53-73.
177. Brosig, J., Weimann, J., & Yang, C. L. (2003). The hot versus cold effect in a simple bargaining experiment. *Experimental Economics*, *6*(1), 75-90.
178. Bru, L., Cabrera, S., Capra, C. M., & Gomez, R. (2003). A common pool resource game with sequential decisions and experimental evidence. *Experimental Economics*, *6*(1), 91-114.
179. Dorsey, R., & Razzolini, L. (2003). Explaining overbidding in first price auctions using controlled lotteries. *Experimental Economics*, *6*(2), 123-140.
180. Grosskopf, B. (2003). Reinforcement and directional learning in the ultimatum game with responder competition. *Experimental Economics*, *6*(2), 141-158.
181. Sgroi, D. (2003). The right choice at the right time: A herding experiment in endogenous time. *Experimental Economics*, *6*(2), 159-180.
182. Cooper, D. J., Feltovich, N., Roth, A. E., & Zwick, R. (2003). Relative versus absolute speed of adjustment in strategic environments: responder behavior in ultimatum games. *Experimental Economics*, *6*(2), 181-207.
183. Davis, D. D., Reilly, R. J., & Wilson, B. J. (2003). Cost structures and Nash play in repeated Cournot games. *Experimental Economics*, *6*(2), 209-226.
184. Palacios-Huerta, I. (2003). Learning to open Monty Hall's doors. *Experimental Economics*, *6*(3), 235-251.
185. Meidinger, C., Rullière, J. L., & Villeval, M. C. (2003). Does team-based compensation give rise to problems when agents vary in their ability?. *Experimental Economics*, *6*(3), 253-272.
186. Nirel, R., & Gorfine, M. (2003). Nonparametric Analysis of Longitudinal Binary Data: An Application to the Intergroup Prisoner's Dilemma Game. *Experimental Economics*, *6*(3), 327-341.
187. Weber, R. A., Camerer, C. F., & Knez, M. (2004). Timing and virtual observability in ultimatum bargaining and "weak link" coordination games. *Experimental Economics*, *7*(1), 25-48.

188. Schmitt, P. M. (2004). On perceptions of fairness: The role of valuations, outside options, and information in ultimatum bargaining games. *Experimental Economics*, *7*(1), 49-73.
189. Stevens, D. E., & Williams, A. W. (2004). Inefficiency in earnings forecasts: Experimental evidence of reactions to positive vs. negative information. *Experimental Economics*, *7*(1), 75-92.
190. Parkhurst, G. M., Shogren, J. F., & Bastian, C. (2004). Repetition, communication, and coordination failure. *Experimental Economics*, *7*(2), 141-152.
191. Levati, M. V., & Neugebauer, T. (2004). An application of the English clock market mechanism to public goods games. *Experimental Economics*, *7*(2), 153-169.
192. Charness, G., Frechette, G. R., & Kagel, J. H. (2004). How robust is laboratory gift exchange?. *Experimental Economics*, *7*(2), 189-205.
193. Dale, D. J. (2004). Charitable lottery structure and fund raising: Theory and evidence. *Experimental Economics*, *7*(3), 217-234.
194. Walker, J. M., & Halloran, M. A. (2004). Rewards and sanctions and the provision of public goods in one-shot settings. *Experimental Economics*, *7*(3), 235-247.
195. Rassenti, S. J., & Wilson, B. J. (2004). How applicable is the dominant firm model of price leadership?. *Experimental Economics*, *7*(3), 271-288.
196. Selten, R., Abbink, K., & Cox, R. (2005). Learning direction theory and the winner's curse. *Experimental Economics*, *8*(1), 5-20.
197. Kachelmeier, S. J., & Towry, K. L. (2005). The limitations of experimental design: A case study involving monetary incentive effects in laboratory markets. *Experimental Economics*, *8*(1), 21-33.
198. Burlando, R. M., & Guala, F. (2005). Heterogeneous agents in public goods experiments. *Experimental Economics*, *8*(1), 35-54.
199. Slonim, R. L. (2005). Competing against experienced and inexperienced players. *Experimental Economics*, *8*(1), 55-75.
200. Davis, D. D., Millner, E. L., & Reilly, R. J. (2005). Subsidy schemes and charitable contributions: A closer look. *Experimental Economics*, *8*(2), 85-106.
201. Seale, D. A., Parco, J. E., Stein, W. E., & Rapoport, A. (2005). Joining a queue or staying out: Effects of information structure and service time on arrival and staying out decisions. *Experimental Economics*, *8*(2), 117-144.
202. Dickinson, D. L. (2005). Bargaining outcomes with double-offer arbitration. *Experimental Economics*, *8*(2), 145-166.
203. kan Holm, H., & Engseld, P. (2005). Choosing bargaining partners—An experimental study on the impact of information about income, status and gender. *Experimental Economics*, *8*(3), 183-216.
204. Hey, J. D., & Lee, J. (2005). Do subjects separate (or are they sophisticated)?. *Experimental Economics*, *8*(3), 233-265.
205. Lindqvist, T., & Stennek, J. (2005). The insiders' dilemma: an experiment on merger formation. *Experimental Economics*, *8*(3), 267-284.
206. Ackert, L. F., Charupat, N., Church, B. K., & Deaves, R. (2006). An experimental examination of the house money effect in a multi-period setting. *Experimental Economics*, *9*(1), 5-16.
207. Fischer, S., Güth, W., Müller, W., & Stiehler, A. (2006). From ultimatum to Nash bargaining: Theory and experimental evidence. *Experimental Economics*, *9*(1), 17-33.
208. Bosman, R., Hennig-Schmidt, H., & Van Winden, F. (2006). Exploring group decision making in a power-to-take experiment. *Experimental Economics*, *9*(1), 35-51.

209. Eckel, C. C., & Wilson, R. K. (2006). Internet cautions: Experimental games with internet partners. *Experimental Economics*, *9*(1), 53-66.
210. Abbink, K., & Hennig-Schmidt, H. (2006). Neutral versus loaded instructions in a bribery experiment. *Experimental Economics*, *9*(2), 103-121.
211. Schotter, A., & Sopher, B. (2006). Trust and trustworthiness in games: An experimental study of intergenerational advice. *Experimental Economics*, *9*(2), 123-145.
212. Murphy, R. O., Rapoport, A., & Parco, J. E. (2006). The breakdown of cooperation in iterative real-time trust dilemmas. *Experimental Economics*, *9*(2), 147-166.
213. Ashraf, N., Bohnet, I., & Piankov, N. (2006). Decomposing trust and trustworthiness. *Experimental Economics*, *9*(3), 193-208.
214. Cox, J. C., & Hayne, S. C. (2006). Barking up the right tree: Are small groups rational agents?. *Experimental Economics*, *9*(3), 209-222.
215. Anderson, L. R., Mellor, J. M., & Milyo, J. (2006). Induced heterogeneity in trust experiments. *Experimental Economics*, *9*(3), 223-235.
216. Berninghaus, S. K., Ehrhart, K. M., & Ott, M. (2006). A network experiment in continuous time: The influence of link costs. *Experimental Economics*, *9*(3), 237-251.
217. Cinyabuguma, M., Page, T., & Putterman, L. (2006). Can second-order punishment deter perverse punishment?. *Experimental Economics*, *9*(3), 265-279.
218. Loewenstein, G., Moore, D. A., & Weber, R. A. (2006). Misperceiving the value of information in predicting the performance of others. *Experimental Economics*, *9*(3), 281-295.
219. McKinney, C. N., & Van Huyck, J. B. (2006). Does seeing more deeply into a game increase one's chances of winning?. *Experimental economics*, *9*(3), 297-303.
220. Cadsby, C. B., Maynes, E., & Trivedi, V. U. (2006). Tax compliance and obedience to authority at home and in the lab: A new experimental approach. *Experimental economics*, *9*(4), 343-359.
221. Bodoff, D., Levecq, H., & Zhang, H. (2006). EDGAR on the internet: The welfare effects of wider information distribution in an experimental market for risky assets. *Experimental Economics*, *9*(4), 361-381.
222. Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2006). Elicitation using multiple price list formats. *Experimental Economics*, *9*(4), 383-405.
223. Brandts, J., & Cooper, D. J. (2006). Observability and overcoming coordination failure in organizations: An experimental study. *Experimental Economics*, *9*(4), 407-423.
224. Yang, C. L., Yue, C. S. J., & Yu, I. T. (2007). The rise of cooperation in correlated matching prisoners dilemma: An experiment. *Experimental Economics*, *10*(1), 3-20.
225. Bullock, D. S., & Rutström, E. E. (2007). Policy making and rent-dissipation: An experimental test. *Experimental Economics*, *10*(1), 21-36.
226. Turocy, T. L., Watson, E., & Battalio, R. C. (2007). Framing the first-price auction. *Experimental Economics*, *10*(1), 37-51.
227. Albert, M., Güth, W., Kirchler, E., & Maciejovsky, B. (2007). Are we nice (r) to nice (r) people?—an experimental analysis. *Experimental Economics*, *10*(1), 53-69.
228. Friedman, D., Pommerenke, K., Lukose, R., Milam, G., & Huberman, B. A. (2007). Searching for the sunk cost fallacy. *Experimental Economics*, *10*(1), 79-104.
229. James, D. (2007). Stability of risk preference parameter estimates within the Becker-DeGroot-Marschak procedure. *Experimental Economics*, *10*(2), 123-141.

230. Healy, P. J., Ledyard, J. O., Noussair, C., Thronson, H., Ulrich, P., & Varsi, G. (2007). Contracting inside an organization: An experimental study. *Experimental Economics*, *10*(2), 143-167.
231. Van Huyck, J. B., Battalio, R. C., & Rankin, F. W. (2007). Evidence on learning in coordination games. *Experimental Economics*, *10*(3), 205-220.
232. Cabrales, A., Nagel, R., & Armenter, R. (2007). Equilibrium selection through incomplete information in coordination games: an experimental study. *Experimental Economics*, *10*(3), 221-234.
233. Hess, R. O., Holt, C. A., & Smith, A. M. (2007). Coordination of strategic responses to security threats: Laboratory evidence. *Experimental Economics*, *10*(3), 235-250.
234. Cason, T. N., & Mui, V. L. (2007). Communication and coordination in the laboratory collective resistance game. *Experimental Economics*, *10*(3), 251-267.
235. Brandts, J., Cooper, D. J., & Fatas, E. (2007). Leadership and overcoming coordination failure with asymmetric costs. *Experimental Economics*, *10*(3), 269-284.
236. Hamman, J., Rick, S., & Weber, R. A. (2007). Solving coordination failure with "all-or-none" group-level incentives. *Experimental Economics*, *10*(3), 285-303.
237. Carare, O., Haruvy, E., & Prasad, A. (2007). Hierarchical thinking and learning in rank order contests. *Experimental Economics*, *10*(3), 305-316.
238. Eckel, C. C., & Wilson, R. K. (2007). Social learning in coordination games: does status matter?. *Experimental Economics*, *10*(3), 317-329.
239. Vandegrift, D., Yavas, A., & Brown, P. M. (2007). Incentive effects and overcrowding in tournaments: An experimental analysis. *Experimental Economics*, *10*(4), 345-368.
240. Poulsen, A. U., & Tan, J. H. (2007). Information acquisition in the ultimatum game: An experimental study. *Experimental Economics*, *10*(4), 391-409.
241. Kroll, S., Cherry, T. L., & Shogren, J. F. (2007). The impact of endowment heterogeneity and origin on contributions in best-shot public good games. *Experimental Economics*, *10*(4), 411-428.
242. Deck, C., Farmer, A., & Zeng, D. Z. (2007). Amended final-offer arbitration over an uncertain value: A comparison with CA and FOA. *Experimental Economics*, *10*(4), 439-454.
243. Cox, J. C., Sadiraj, K., & Sadiraj, V. (2008). Implications of trust, fear, and reciprocity for modeling economic behavior. *Experimental Economics*, *11*(1), 1-24.
244. Vyrastekova, J., & Van Soest, D. (2008). On the (in) effectiveness of rewards in sustaining cooperation. *Experimental Economics*, *11*(1), 53-65.
245. Goswami, G., Grace, M. F., & Rebello, M. J. (2008). Experimental evidence on coverage choices and contract prices in the market for corporate insurance. *Experimental Economics*, *11*(1), 67-95.
246. Sbriglia, P. (2008). Revealing the depth of reasoning in p-beauty contest games. *Experimental Economics*, *11*(2), 107-121.
247. Bardsley, N. (2008). Dictator game giving: altruism or artefact?. *Experimental Economics*, *11*(2), 122-133.
248. Iyengar, R., & Schotter, A. (2008). Learning under supervision: an experimental study. *Experimental Economics*, *11*(2), 154-173.
249. Karni, E., Salmon, T., & Sopher, B. (2008). Individual sense of fairness: an experimental study. *Experimental Economics*, *11*(2), 174-189.
250. Neugebauer, T., & Perote, J. (2008). Bidding 'as if'risk neutral in experimental first price auctions without information feedback. *Experimental Economics*, *11*(2), 190-202.

251. Dasgupta, S., Randazzo, K. A., Sheehan, R. S., & Williams, K. C. (2008). Coordinated voting in sequential and simultaneous elections: some experimental evidence. *Experimental Economics*, *11*(4), 315-335.
252. Yamamori, T., Kato, K., Kawagoe, T., & Matsui, A. (2008). Voice matters in a dictator game. *Experimental Economics*, *11*(4), 336-343.
253. Katok, E., & Kwasnica, A. M. (2008). Time is money: The effect of clock speed on seller's revenue in Dutch auctions. *Experimental Economics*, *11*(4), 344-357.
254. Nikiforakis, N., & Normann, H. T. (2008). A comparative statics analysis of punishment in public-good experiments. *Experimental Economics*, *11*(4), 358-369.
255. Kurzban, R., Rigdon, M. L., & Wilson, B. J. (2008). Incremental approaches to establishing trust. *Experimental Economics*, *11*(4), 370-389.
256. Orzen, H. (2008). Counterintuitive number effects in experimental oligopolies. *Experimental Economics*, *11*(4), 390-401.
257. Grimm, V., Kovarik, J., & Ponti, G. (2008). Fixed price plus rationing: an experiment. *Experimental Economics*, *11*(4), 402-422.
258. Deck, C. (2009). An experimental analysis of cooperation and productivity in the trust game. *Experimental economics*, *12*(1), 1-11.
259. Bone, J., Hey, J. D., & Suckling, J. (2009). Do people plan?. *Experimental Economics*, *12*(1), 12-25.
260. Luhan, W. J., Kocher, M. G., & Sutter, M. (2009). Group polarization in the team dictator game reconsidered. *Experimental Economics*, *12*(1), 26-41.
261. Sherstyuk, K. (2009). A comparison of first price multi-object auctions. *Experimental Economics*, *12*(1), 42-64.
262. Davis, D., Korenok, O., & Reilly, R. (2009). Re-matching, information and sequencing effects in posted offer markets. *Experimental Economics*, *12*(1), 65-86.
263. Herrmann, B., & Thöni, C. (2009). Measuring conditional cooperation: a replication study in Russia. *Experimental Economics*, *12*(1), 87-92.
264. Aguiar, F., Brañas-Garza, P., Cobo-Reyes, R., Jimenez, N., & Miller, L. M. (2009). Are women expected to be more generous?. *Experimental Economics*, *12*(1), 93-98.
265. Harrison, G. W., & Rutström, E. E. (2009). Expected utility theory and prospect theory: One wedding and a decent funeral. *Experimental economics*, *12*(2), 133-158.
266. Chou, E., McConnell, M., Nagel, R., & Plott, C. R. (2009). The control of game form recognition in experiments: Understanding dominant strategy failures in a simple two person "guessing" game. *Experimental Economics*, *12*(2), 159-179.
267. Hurkens, S., & Kartik, N. (2009). Would I lie to you? On social preferences and lying aversion. *Experimental Economics*, *12*(2), 180-192.
268. Zhang, P. (2009). Uniform price auctions and fixed price offerings in IPOs: an experimental comparison. *Experimental Economics*, *12*(2), 202-219.
269. Sánchez-Pagés, S., & Vorsatz, M. (2009). Enjoy the silence: an experiment on truth-telling. *Experimental Economics*, *12*(2), 220-241.
270. Stahl, D. O., & Haruvy, E. (2009). Testing theories of behavior for extensive-form two-player two-stage games. *Experimental economics*, *12*(2), 242-251.
271. Goerg, S. J., & Selten, R. (2009). Experimental investigation of stationary concepts in cyclic duopoly games. *Experimental Economics*, *12*(3), 253-271.
272. Carpenter, J., & Matthews, P. H. (2009). What norms trigger punishment?. *Experimental Economics*, *12*(3), 272-288.
273. Mago, S. D., & Dechenaux, E. (2009). Price leadership and firm size asymmetry: an experimental analysis. *Experimental Economics*, *12*(3), 289-317.
274. Sutter, M., Bosman, R., Kocher, M. G., & van Winden, F. (2009). Gender pairing and bargaining—Beware the same sex!. *Experimental Economics*, *12*(3), 318-331.

275. Blavatskyy, P. R., & Köhler, W. R. (2009). Range effects and lottery pricing. *Experimental Economics*, *12*(3), 332-349.
276. Becker, O., Leitner, J., & Leopold-Wildburger, U. (2009). Expectation formation and regime switches. *Experimental Economics*, *12*(3), 350-364.
277. Bruner, D. M. (2009). Changing the probability versus changing the reward. *Experimental Economics*, *12*(4), 367-385.
278. Healy, A. (2009). How effectively do people learn from a variety of different opinions?. *Experimental Economics*, *12*(4), 386-416.
279. Boone, J., Chen, R., Goeree, J. K., & Polydoro, A. (2009). Risky procurement with an insider bidder. *Experimental Economics*, *12*(4), 417-436.
280. Becker, A., & Miller, L. M. (2009). Promoting justice by treating people unequally: an experimental study. *Experimental Economics*, *12*(4), 437-449.
281. Hyndman, K., Terracol, A., & Vaksmann, J. (2009). Learning and sophistication in coordination games. *Experimental Economics*, *12*(4), 450-472.
282. Rydval, O., Ortmann, A., Prokosheva, S., & Hertwig, R. (2009). How certain is the uncertainty effect?. *Experimental Economics*, *12*(4), 473-487.
283. Barr, A., & Serra, D. (2009). The effects of externalities and framing on bribery in a petty corruption experiment. *Experimental Economics*, *12*(4), 488-503.
284. Tufano, F. (2010). Are 'true'preferences revealed in repeated markets? An experimental demonstration of context-dependent valuations. *Experimental Economics*, *13*(1), 1-13.
285. Isaac, R. M., Pevnitskaya, S., & Salmon, T. C. (2010). Do preferences for charitable giving help auctioneers?. *Experimental economics*, *13*(1), 14-44.
286. Davis, D., Korenok, O., & Reilly, R. (2010). Cooperation without coordination: Signaling, types and tacit collusion in laboratory oligopolies. *Experimental economics*, *13*(1), 45-65.
287. Cox, J. C., Servátka, M., & Vadovič, R. (2010). Saliency of outside options in the lost wallet game. *Experimental Economics*, *13*(1), 66-74.
288. Ziegelmeyer, A., Koessler, F., Bracht, J., & Winter, E. (2010). Fragility of information cascades: an experimental study using elicited beliefs. *Experimental Economics*, *13*(2), 121-145.
289. Engle-Warnick, J., & Turdaliev, N. (2010). An experimental test of Taylor-type rules with inexperienced central bankers. *Experimental Economics*, *13*(2), 146-166.
290. Kagel, J. H., Sung, H., & Winter, E. (2010). Veto power in committees: an experimental study. *Experimental Economics*, *13*(2), 167-188.
291. Davis, L. R., Joyce, B. P., & Roelofs, M. R. (2010). My money or yours: house money payment effects. *Experimental Economics*, *13*(2), 189-205.
292. Poulsen, A. U., & Roos, M. W. (2010). Do people make strategic commitments? Experimental evidence on strategic information avoidance. *Experimental Economics*, *13*(2), 206-225.
293. Coricelli, G., Joffily, M., Montmarquette, C., & Villeval, M. C. (2010). Cheating, emotions, and rationality: an experiment on tax evasion. *Experimental Economics*, *13*(2), 226-247.
294. Engelmann, D., & Normann, H. T. (2010). Maximum effort in the minimum-effort game. *Experimental Economics*, *13*(3), 249-259.
295. Bornhorst, F., Ichino, A., Kirchkamp, O., Schlag, K. H., & Winter, E. (2010). Similarities and differences when building trust: the role of cultures. *Experimental Economics*, *13*(3), 260-283.

296. Cadsby, C. B., Servátka, M., & Song, F. (2010). Gender and generosity: does degree of anonymity or group gender composition matter?. *Experimental economics*, *13*(3), 299-308.
297. Bellemare, C., Bissonnette, L., & Kröger, S. (2010). Bounding preference parameters under different assumptions about beliefs: a partial identification approach. *Experimental economics*, *13*(3), 334-345.
298. Chaudhuri, A., & Paichayontvijit, T. (2010). Recommended play and performance bonuses in the minimum effort coordination game. *Experimental Economics*, *13*(3), 346-363.
299. Gächter, S., & Renner, E. (2010). The effects of (incentivized) belief elicitation in public goods experiments. *Experimental Economics*, *13*(3), 364-377.
300. Veiga, H., & Vorsatz, M. (2010). Information aggregation in experimental asset markets in the presence of a manipulator. *Experimental Economics*, *13*(4), 379-398.
301. Fatas, E., Meléndez-Jiménez, M. A., & Solaz, H. (2010). An experimental analysis of team production in networks. *Experimental Economics*, *13*(4), 399-411.
302. Blanco, M., Engelmann, D., Koch, A. K., & Normann, H. T. (2010). Belief elicitation in experiments: is there a hedging problem?. *Experimental Economics*, *13*(4), 412-438.
303. Darai, D., Sacco, D., & Schmutzler, A. (2010). Competition and innovation: An experimental investigation. *Experimental Economics*, *13*(4), 439-460.
304. Takeuchi, K., Lin, J. C., Chen, Y., & Finholt, T. A. (2010). Scheduling with package auctions. *Experimental economics*, *13*(4), 476-499.
305. Krawczyk, M., & Le Lec, F. (2010). 'Give me a chance!'An experiment in social decision under risk. *Experimental Economics*, *13*(4), 500-511.
306. Houser, D., & Xiao, E. (2011). Classification of natural language messages using a coordination game. *Experimental Economics*, *14*(1), 1-14.
307. Kimbrough, E. O. (2011). Learning to respect property by refashioning theft into trade. *Experimental economics*, *14*(1), 84-109.
308. Comerton-Forde, C., & Putniņš, T. J. (2011). Pricing accuracy, liquidity and trader behavior with closing price manipulation. *Experimental economics*, *14*(1), 110-131.
309. Cabrales, A., Charness, G., & Villeval, M. C. (2011). Hidden information, bargaining power, and efficiency: an experiment. *Experimental Economics*, *14*(2), 133-159.
310. Iriberri, N., & Rey-Biel, P. (2011). The role of role uncertainty in modified dictator games. *Experimental Economics*, *14*(2), 160-180.
311. Mittone, L., & Ploner, M. (2011). Peer pressure, social spillovers, and reciprocity: an experimental analysis. *Experimental Economics*, *14*(2), 203-222.
312. Huber, J., Angerer, M., & Kirchler, M. (2011). Experimental asset markets with endogenous choice of costly asymmetric information. *Experimental Economics*, *14*(2), 223-240.
313. Cubitt, R. P., Drouvelis, M., & Gächter, S. (2011). Framing and free riding: emotional responses and punishment in social dilemma games. *Experimental Economics*, *14*(2), 254-272.
314. Gazzale, R. S., & Khopkar, T. (2011). Remain silent and ye shall suffer: seller exploitation of reticent buyers in an experimental reputation system. *Experimental Economics*, *14*(2), 273-285.
315. Uler, N. (2011). Public goods provision, inequality and taxes. *Experimental Economics*, *14*(3), 287-306.
316. Ackert, L. F., Gillette, A. B., Martinez-Vazquez, J., & Rider, M. (2011). Are benevolent dictators altruistic in groups? A within-subject design. *Experimental Economics*, *14*(3), 307-321.

317. Otto, P. E., & Bolle, F. (2011). Matching markets with price bargaining. *Experimental Economics*, *14*(3), 322-348.
318. Ballinger, T. P., Hudson, E., Karkoviata, L., & Wilcox, N. T. (2011). Saving behavior and cognitive abilities. *Experimental Economics*, *14*(3), 349-374.
319. Anderson, L. R., DiTraglia, F. J., & Gerlach, J. R. (2011). Measuring altruism in a public goods experiment: a comparison of US and Czech subjects. *Experimental Economics*, *14*(3), 426-437.
320. Osés-Eraso, N., & Viladrich-Grau, M. (2011). The sustainability of the commons: giving and receiving. *Experimental Economics*, *14*(4), 458-481.
321. Visser, M. S., & Roelofs, M. R. (2011). Heterogeneous preferences for altruism: gender and personality, social status, giving and taking. *Experimental Economics*, *14*(4), 490-506.
322. Vieider, F. M. (2011). Separating real incentives and accountability. *Experimental Economics*, *14*(4), 507-518.
323. Janssen, M. A., Anderies, J. M., & Joshi, S. R. (2011). Coordination and cooperation in asymmetric commons dilemmas. *Experimental Economics*, *14*(4), 547-566.
324. Di Mauro, C., & Castro, M. F. (2011). Kindness, confusion, or… ambiguity?. *Experimental Economics*, *14*(4), 611-633.
325. Proto, E., Sgroi, D., & Oswald, A. J. (2012). Are happiness and productivity lower among young people with newly-divorced parents? An experimental and econometric approach. *Experimental Economics*, *15*(1), 1-23.
326. Reuben, E., & Suetens, S. (2012). Revisiting strategic versus non-strategic cooperation. *Experimental Economics*, *15*(1), 24-43.
327. Deck, C., & Nikiforakis, N. (2012). Perfect and imperfect real-time monitoring in a minimum-effort game. *Experimental Economics*, *15*(1), 71-88.
328. Huber, J., & Kirchler, M. (2012). The impact of instructions and procedure on reducing confusion and bubbles in experimental asset markets. *Experimental Economics*, *15*(1), 89-105.
329. Goertz, J. M. (2012). Market composition and experience in common-value auctions. *Experimental Economics*, *15*(1), 106-127.
330. Lévy-Garboua, L., Maafi, H., Masclet, D., & Terracol, A. (2012). Risk aversion and framing effects. *Experimental Economics*, *15*(1), 128-144.
331. Rosenboim, M., & Shavit, T. (2012). Whose money is it anyway? Using prepaid incentives in experimental economics to create a natural environment. *Experimental Economics*, *15*(1), 145-157.
332. Rodriguez-Lara, I., & Moreno-Garrido, L. (2012). Self-interest and fairness: self-serving choices of justice principles. *Experimental Economics*, *15*(1), 158-175.
333. Hyndman, K., Özbay, E. Y., Schotter, A., & Ehrblatt, W. (2012). Belief formation: an experiment with outside observers. *Experimental Economics*, *15*(1), 176-203.
334. Keck, S., & Karelaia, N. (2012). Does competition foster trust? The role of tournament incentives. *Experimental Economics*, *15*(1), 204-228.
335. Reinstein, D., & Riener, G. (2012). Decomposing desert and tangibility effects in a charitable giving experiment. *Experimental Economics*, *15*(1), 229-240.
336. Casari, M., & Luini, L. (2012). Peer punishment in teams: expressive or instrumental choice?. *Experimental Economics*, *15*(2), 241-259.
337. Norton, D. A., & Isaac, R. M. (2012). Experts with a conflict of interest: a source of ambiguity?. *Experimental economics*, *15*(2), 260-277.
338. Cabrales, A., Nagel, R., & Mora, J. V. R. (2012). It is Hobbes, not Rousseau: an experiment on voting and redistribution. *Experimental Economics*, *15*(2), 278-308.

339. Cramton, P., Filiz-Ozbay, E., Ozbay, E. Y., & Sujarittanonta, P. (2012). Discrete clock auctions: an experimental study. *Experimental Economics*, *15*(2), 309-322.
340. Ziegelmeyer, A., Schmelz, K., & Ploner, M. (2012). Hidden costs of control: four repetitions and an extension. *Experimental Economics*, *15*(2), 323-340.
341. Cheung, S. L., & Palan, S. (2012). Two heads are less bubbly than one: team decision-making in an experimental asset market. *Experimental Economics*, *15*(3), 373-397.
342. Greiner, B., & Güth, W. (2012). Social communication and discrimination: a video experiment. *Experimental Economics*, *15*(3), 398-417.
343. Baltussen, G., Post, G. T., Van Den Assem, M. J., & Wakker, P. P. (2012). Random incentive systems in a dynamic choice experiment. *Experimental Economics*, *15*(3), 418-443.
344. Dickinson, D. L., & McElroy, T. (2012). Circadian effects on strategic reasoning. *Experimental Economics*, *15*(3), 444-459.
345. Dannenberg, A., Riechmann, T., Sturm, B., & Vogt, C. (2012). Inequality aversion and the house money effect. *Experimental Economics*, *15*(3), 460-484.
346. Gürerk, Ö., & Selten, R. (2012). The effect of payoff tables on experimental oligopoly behavior. *Experimental Economics*, *15*(3), 499-509.
347. Pearson, M., & Schipper, B. C. (2012). The visible hand: finger ratio (2D: 4D) and competitive bidding. *Experimental Economics*, *15*(3), 510-529.
348. Cobo-Reyes, R., & Jiménez, N. (2012). The dark side of friendship:'envy'. *Experimental Economics*, *15*(4), 547-570.
349. Aimone, J. A., & Houser, D. (2012). What you don't know won't hurt you: a laboratory analysis of betrayal aversion. *Experimental Economics*, *15*(4), 571-588.
350. Ong, Q., Riyanto, Y. E., & Sheffrin, S. M. (2012). How does voice matter? Evidence from the ultimatum game. *Experimental Economics*, *15*(4), 604-621.
351. Danz, D. N., Fehr, D., & Kübler, D. (2012). Information and beliefs in a repeated normal-form game. *Experimental Economics*, *15*(4), 622-640.
352. Buser, T., & Peter, N. (2012). Multitasking. *Experimental Economics*, *15*(4), 641-655.
353. Scheffel, T., Ziegler, G., & Bichler, M. (2012). On the impact of package selection in combinatorial auctions: an experimental study in the context of spectrum auction design. *Experimental Economics*, *15*(4), 667-692.
354. Levati, M. V., Qiu, J., & Mahagaonkar, P. (2012). Testing the Modigliani-Miller theorem directly in the lab. *Experimental Economics*, *15*(4), 693-716.
355. Klijn, F., Pais, J., & Vorsatz, M. (2013). Preference intensities and risk aversion in school choice: A laboratory experiment. *Experimental Economics*, *16*(1), 1-22.
356. Feri, F., Gantner, A., Höchtl, W., & Sausgruber, R. (2013). The pivotal mechanism revisited: some evidence on group manipulation. *Experimental Economics*, *16*(1), 23-51.
357. Goeree, J. K., Offerman, T., & Sloof, R. (2013). Demand reduction and preemptive bidding in multi-unit license auctions. *Experimental Economics*, *16*(1), 52-87.
358. Hortala-Vallve, R., Llorente-Saguer, A., & Nagel, R. (2013). The role of information in different bargaining protocols. *Experimental Economics*, *16*(1), 88-113.
359. Sherstyuk, K., Tarui, N., & Saijo, T. (2013). Payment schemes in infinite-horizon experimental games. *Experimental Economics*, *16*(1), 125-153.
360. Jakiela, P. (2013). Equity vs. efficiency vs. self-interest: on the use of dictator games to measure distributional preferences. *Experimental Economics*, *16*(2), 208-221.

361. López-Pérez, R., & Spiegelman, E. (2013). Why do people tell the truth? Experimental evidence for pure lie aversion. *Experimental Economics*, *16*(3), 233-247.
362. Li, K. K. (2013). Asymmetric memory recall of positive and negative events in social interactions. *Experimental Economics*, *16*(3), 248-262.
363. Filippin, A., & Guala, F. (2013). Costless discrimination and unequal achievements in an experimental tournament. *Experimental economics*, *16*(3), 285-305.
364. Oexl, R., & Grossman, Z. J. (2013). Shifting the blame to a powerless intermediary. *Experimental Economics*, *16*(3), 306-312.
365. Shurchkov, O. (2013). Coordination and learning in dynamic global games: experimental evidence. *Experimental Economics*, *16*(3), 313-334.
366. Cleave, B. L., Nikiforakis, N., & Slonim, R. (2013). Is there selection bias in laboratory experiments? The case of social and risk preferences. *Experimental Economics*, *16*(3), 372-382.
367. Smith, A. (2013). Estimating the causal effect of beliefs on contributions in repeated public good games. *Experimental Economics*, *16*(3), 414-425.
368. Cabrera, S., Fatás, E., Lacomba, J. A., & Neugebauer, T. (2013). Splitting leagues: promotion and demotion in contribution-based regrouping experiments. *Experimental Economics*, *16*(3), 426-441.
369. Cooper, D. J., & Lightle, J. P. (2013). The gift of advice: Communication in a bilateral gift exchange game. *Experimental Economics*, *16*(4), 443-477.
370. Bayer, R. C., Renner, E., & Sausgruber, R. (2013). Confusion and learning in the voluntary contributions game. *Experimental Economics*, *16*(4), 478-496.
371. Cherry, T. L., Cotten, S. J., & Kroll, S. (2013). Heterogeneity, coordination and the provision of best-shot public goods. *Experimental Economics*, *16*(4), 497-510.
372. Bichler, M., Shabalin, P., & Wolf, J. (2013). Do core-selecting combinatorial clock auctions always lead to high efficiency? An experimental analysis of spectrum auction designs. *Experimental Economics*, *16*(4), 511-545.
373. Anbarci, N., & Feltovich, N. (2013). How sensitive are bargaining outcomes to changes in disagreement payoffs?. *Experimental Economics*, *16*(4), 560-596.
374. Deck, C., Servátka, M., & Tucker, S. (2013). An examination of the effect of messages on cooperation under double-blind and single-blind payoff procedures. *Experimental Economics*, *16*(4), 597-607.
375. Nikiforakis, N., & Mitchell, H. (2014). Mixing the carrots with the sticks: Third party punishment and reward. *Experimental Economics*, *17*(1), 1-23.
376. Cárdenas, J. C., De Roux, N., Jaramillo, C. R., & Martinez, L. R. (2014). Is it my money or not? An experiment on risk aversion and the house-money effect. *Experimental Economics*, *17*(1), 47-60.
377. Cornand, C., & Heinemann, F. (2014). Measuring agents' reaction to private and public information in games with strategic complementarities. *Experimental Economics*, *17*(1), 61-77.
378. Drugov, M., Hamman, J., & Serra, D. (2014). Intermediaries in corruption: an experiment. *Experimental Economics*, *17*(1), 78-99.
379. Aycinena, D., Baltaduonis, R., & Rentschler, L. (2014). Valuation structure in first-price and least-revenue auctions: an experimental investigation. *Experimental Economics*, *17*(1), 100-128.
380. Cheung, S. L. (2014). New insights into conditional cooperation and punishment from a strategy method experiment. *Experimental Economics*, *17*(1), 129-153.

381. de Groot Ruiz, A., Offerman, T., & Onderstal, S. (2014). For those about to talk we salute you: an experimental study of credible deviations and ACDC. *Experimental Economics*, *17*(2), 173-199.
382. Filiz-Ozbay, E., & Ozbay, E. Y. (2014). Effect of an audience in public goods provision. *Experimental Economics*, *17*(2), 200-214.
383. Tsutsui, K., & Zizzo, D. J. (2014). Group status, minorities and trust. *Experimental Economics*, *17*(2), 215-244.
384. Brookins, P., & Ryvkin, D. (2014). An experimental study of bidding in contests of incomplete information. *Experimental Economics*, *17*(2), 245-261.
385. Mak, V., & Zwick, R. (2014). Experimenting and learning with localized direct communication. *Experimental Economics*, *17*(2), 262-284.
386. Zhang, B., Li, C., De Silva, H., Bednarik, P., & Sigmund, K. (2014). The evolution of sanctioning institutions: an experimental approach to the social contract. *Experimental Economics*, *17*(2), 285-303.
387. Stöckl, T. (2014). Price efficiency and trading behavior in limit order markets with competing insiders. *Experimental Economics*, *17*(2), 314-334.
388. Chuah, S. H., Hoffmann, R., & Larner, J. (2014). Elicitation effects in a multi-stage bargaining experiment. *Experimental Economics*, *17*(2), 335-345.
389. Koppel, H., & Regner, T. (2014). Corporate Social Responsibility in the work place. *Experimental Economics*, *17*(3), 347-370.
390. Ruud, P. A., Schunk, D., & Winter, J. K. (2014). Uncertainty causes rounding: an experimental study. *Experimental Economics*, *17*(3), 391-413.
391. Goeree, J. K., & Zhang, J. (2014). Communication & competition. *Experimental Economics*, *17*(3), 421-438.
392. Jacobson, S., & Petrie, R. (2014). Favor trading in public good provision. *Experimental Economics*, *17*(3), 439-460.
393. Dechenaux, E., Mago, S. D., & Razzolini, L. (2014). Traffic congestion: an experimental study of the Downs-Thomson paradox. *Experimental economics*, *17*(3), 461-487.
394. Korenok, O., Millner, E. L., & Razzolini, L. (2014). Taking, giving, and impure altruism in dictator games. *Experimental Economics*, *17*(3), 488-500.
395. Chen, Y., Jeon, G. Y., & Kim, Y. M. (2014). A day without a search engine: An experimental study of online and offline searches. *Experimental Economics*, *17*(4), 512-536.
396. Butler, D., Isoni, A., Loomes, G., & Tsutsui, K. (2014). Beyond choice: investigating the sensitivity and validity of measures of strength of preference. *Experimental Economics*, *17*(4), 537-563.
397. Agranov, M., Bisin, A., & Schotter, A. (2014). An experimental study of the impact of competition for other people's money: the portfolio manager market. *Experimental Economics*, *17*(4), 564-585.
398. Reiß, J. P., & Wolff, I. (2014). Incentive effects of funding contracts: an experiment. *Experimental Economics*, *17*(4), 586-614.
399. Kvaløy, O., & Luzuriaga, M. (2014). Playing the trust game with other people's money. *Experimental Economics*, *17*(4), 615-630.
400. Botelho, A., Dinar, A., Pinto, L. M. C., & Rapoport, A. (2014). Time and uncertainty in resource dilemmas: equilibrium solutions and experimental results. *Experimental economics*, *17*(4), 649-672.
401. Samek, A. S., & Sheremeta, R. M. (2014). Recognizing contributors: an experiment on public goods. *Experimental Economics*, *17*(4), 673-690.

402. Buchanan, J. A., & Wilson, B. J. (2014). An experiment on protecting intellectual property. *Experimental Economics*, *17*(4), 691-716.
403. Cojoc, D., & Stoian, A. (2014). Dishonesty and charitable behavior. *Experimental Economics*, *17*(4), 717-732.
404. Nosenzo, D., Quercia, S., & Sefton, M. (2015). Cooperation in small groups: The effect of group size. *Experimental Economics*, *18*(1), 4-14.
405. Leibbrandt, A., Ramalingam, A., Sääksvuori, L., & Walker, J. M. (2015). Incomplete punishment networks in public goods games: experimental evidence. *Experimental Economics*, *18*(1), 15-37.
406. Kamei, K., Putterman, L., & Tyran, J. R. (2015). State or nature? Endogenous formal versus informal sanctions in the voluntary provision of public goods. *Experimental Economics*, *18*(1), 38-65.
407. Cason, T. N., & Gangadharan, L. (2015). Promoting cooperation in nonlinear social dilemmas through peer punishment. *Experimental Economics*, *18*(1), 66-88.
408. Cherry, J., Salant, S., & Uler, N. (2015). Experimental departures from self-interest when competing partnerships share output. *Experimental Economics*, *18*(1), 89-115.
409. de Oliveira, A. C., Croson, R. T., & Eckel, C. (2015). One bad apple? Heterogeneity and information in public good provision. *Experimental Economics*, *18*(1), 116-135.
410. Dickinson, D. L., Dutcher, E. G., & Rodet, C. S. (2015). Observed punishment spillover effects: a laboratory investigation of behavior in a social dilemma. *Experimental Economics*, *18*(1), 136-153.
411. Andreoni, J., & Gee, L. K. (2015). Gunning for efficiency with third party enforcement in threshold public goods. *Experimental Economics*, *18*(1), 154-171.
412. Brown, A. L., & Cohen, G. (2015). Does anonymity affect the willingness to accept and willingness to pay gap? A generalization of Plott and Zeiler. *Experimental Economics*, *18*(2), 173-184.
413. Roux, C., & Thöni, C. (2015). Do control questions influence behavior in experiments?. *Experimental Economics*, *18*(2), 185-194.
414. Cox, J. C., Sadiraj, V., & Schmidt, U. (2015). Paradoxes and mechanisms for choice under risk. *Experimental Economics*, *18*(2), 215-250.
415. Corgnet, B., Hernán-González, R., & Schniter, E. (2015). Why real leisure really matters: Incentive effects on real effort in the laboratory. *Experimental Economics*, *18*(2), 284-301.
416. Stöckl, T., Huber, J., & Kirchler, M. (2015). Multi-period experimental asset markets with distinct fundamental value regimes. *Experimental Economics*, *18*(2), 314-334.
417. Gretschko, V., & Rajko, A. (2015). Excess information acquisition in auctions. *Experimental Economics*, *18*(3), 335-355.
418. Kartal, M. (2015). Laboratory elections with endogenous turnout: proportional representation versus majoritarian rule. *Experimental Economics*, *18*(3), 366-384.
419. Wibral, M. (2015). Identity changes and the efficiency of reputation systems. *Experimental Economics*, *18*(3), 408-431.
420. Khadjavi, M., & Lange, A. (2015). Doing good or doing harm: experimental evidence on giving and taking in public good games. *Experimental Economics*, *18*(3), 432-441.
421. Noussair, C. N., & Stoop, J. (2015). Time as a medium of reward in three social preference experiments. *Experimental Economics*, *18*(3), 442-456.
422. Harris, D., Herrmann, B., Kontoleon, A., & Newton, J. (2015). Is it a norm to favour your own group?. *Experimental Economics*, *18*(3), 491-521.
423. Robbett, A. (2015). Voting with hands and feet: the requirements for optimal group formation. *Experimental Economics*, *18*(3), 522-541.

424. Bonein, A., & Denant-Boèmont, L. (2015). Self-control, commitment and peer pressure: a laboratory experiment. *Experimental Economics*, *18*(4), 543-568.
425. Neri, C. (2015). Eliciting beliefs in continuous-choice games: a double auction experiment. *Experimental Economics*, *18*(4), 569-608.
426. Giamattei, M., & Lambsdorff, J. G. (2015). Balancing the current account: experimental evidence on underconsumption. *Experimental Economics*, *18*(4), 670-696.
427. Tergiman, C. (2015). Institution design and public good provision: an experimental study of the vote of confidence procedure. *Experimental Economics*, *18*(4), 697-717.
428. Fielding, D., & Knowles, S. (2015). Can you spare some change for charity? Experimental evidence on verbal cues and loose change effects in a Dictator Game. *Experimental Economics*, *18*(4), 718-730.
429. Offerman, T., & Palley, A. B. (2016). Lossed in translation: an off-the-shelf method to recover probabilistic beliefs from loss-averse agents. *Experimental economics*, *19*(1), 1-30.
430. Kalaycı, K., & Serra-Garcia, M. (2016). Complexity and biases. *Experimental Economics*, *19*(1), 31-50.
431. Kloosterman, A. (2016). Directed search with heterogeneous firms: an experimental study. *Experimental Economics*, *19*(1), 51-66.
432. Davis, D., Ivanov, A., & Korenok, O. (2016). Individual characteristics and behavior in repeated games: an experimental study. *Experimental Economics*, *19*(1), 67-99.
433. Di Bartolomeo, G., & Papa, S. (2016). Trust and reciprocity: extensions and robustness of triadic design. *Experimental Economics*, *19*(1), 100-115.
434. Lien, J. W., Zheng, J., & Zhong, X. (2016). Preference submission timing in school choice matching: testing fairness and efficiency in the laboratory. *Experimental Economics*, *19*(1), 116-150.
435. Mengel, F., Tsakas, E., & Vostroknutov, A. (2016). Past experience of uncertainty affects risk aversion. *Experimental Economics*, *19*(1), 151-176.
436. Devetag, G., Di Guida, S., & Polonio, L. (2016). An eye-tracking study of feature-based choice in one-shot games. *Experimental Economics*, *19*(1), 177-201.
437. Kamijo, Y., Ozono, H., & Shimizu, K. (2016). Overcoming coordination failure using a mechanism based on gradualism and endogeneity. *Experimental Economics*, *19*(1), 202-217.
438. Thoma, C. (2016). Under-versus overconfidence: an experiment on how others perceive a biased self-assessment. *Experimental Economics*, *19*(1), 218-239.
439. Banerjee, R. (2016). On the interpretation of bribery in a laboratory corruption game: moral frames and social norms. *Experimental Economics*, *19*(1), 240-267.
440. Erat, S., & Gneezy, U. (2016). Incentives for creativity. *Experimental Economics*, *19*(2), 269-280.
441. Meissner, T. (2016). Intertemporal consumption and debt aversion: an experimental study. *Experimental Economics*, *19*(2), 281-298.
442. Kalaycı, K. (2016). Confusopoly: competition and obfuscation in markets. *Experimental Economics*, *19*(2), 299-316.
443. Duffy, J., & Lafky, J. (2016). Birth, death and public good provision. *Experimental Economics*, *19*(2), 317-341.
444. Feldhaus, C., & Stauf, J. (2016). More than words: the effects of cheap talk in a volunteer's dilemma. *Experimental Economics*, *19*(2), 342-359.
445. Tyszler, M., & Schram, A. (2016). Information and strategic voting. *Experimental economics*, *19*(2), 360-381.

446. Ismayilov, H., & Potters, J. (2016). Why do promises affect trustworthiness, or do they?. *Experimental Economics*, *19*(2), 382-393.
447. Huber, J., Kirchler, M., & Stöckl, T. (2016). The influence of investment experience on market prices: laboratory evidence. *Experimental Economics*, *19*(2), 394-411.
448. Charness, G., Cobo-Reyes, R., Lacomba, J. A., Lagos, F., & Pérez, J. M. (2016). Social comparisons in wage delegation: Experimental evidence. *Experimental Economics*, *19*(2), 433-459.
449. Montinari, N., Nicolò, A., & Oexl, R. (2016). The gift of being chosen. *Experimental Economics*, *19*(2), 460-479.
450. Riyanto, Y. E., & Zhang, J. (2016). Putting a price tag on others' perceptions of us. *Experimental Economics*, *19*(2), 480-499.
451. Yamakawa, T., Okano, Y., & Saijo, T. (2016). Detecting motives for cooperation in public goods experiments. *Experimental Economics*, *19*(2), 500-512.
452. Bolton, G., & Werner, P. (2016). The influence of potential on wages and effort. *Experimental Economics*, *19*(3), 535-561.
453. Hauge, K. E., Brekke, K. A., Johansson, L. O., Johansson-Stenman, O., & Svedsäter, H. (2016). Keeping others in our mind or in our heart? Distribution games under cognitive load. *Experimental Economics*, *19*(3), 562-576.
454. Hey, J. D., & Di Cagno, D. (2016). Does money impede convergence?. *Experimental Economics*, *19*(3), 595-612.
455. Crosetto, P., & Filippin, A. (2016). A theoretical and experimental appraisal of four risk elicitation methods. *Experimental Economics*, *19*(3), 613-641.
456. Ifcher, J., & Zarghamee, H. (2016). Pricing competition: a new laboratory measure of gender differences in the willingness to compete. *Experimental Economics*, *19*(3), 642-662.
457. Fehr, D., & Huck, S. (2016). Who knows it is a game? On strategic awareness and cognitive ability. *Experimental Economics*, *19*(4), 713-726.
458. Casari, M., Zhang, J., & Jackson, C. (2016). Same process, different outcomes: group performance in an acquiring a company experiment. *Experimental Economics*, *19*(4), 764-791.
459. Brandts, J., Rott, C., & Solà, C. (2016). Not just like starting over-Leadership and revivification of cooperation in groups. *Experimental economics*, *19*(4), 792-818.
460. Grimalda, G., Kar, A., & Proto, E. (2016). Procedural fairness in lotteries assigning initial roles in a dynamic setting. *Experimental Economics*, *19*(4), 819-841.
461. Frackenpohl, G., Hillenbrand, A., & Kube, S. (2016). Leadership effectiveness and institutional frames. *Experimental Economics*, *19*(4), 842-863.
462. Olivola, C. Y., & Wang, S. W. (2016). Patience auctions: The impact of time vs. money bidding on elicited discount rates. *Experimental Economics*, *19*(4), 864-885.
463. Aguiar-Conraria, L., Magalhães, P. C., & Vanberg, C. A. (2016). Experimental evidence that quorum rules discourage turnout and promote election boycotts. *Experimental Economics*, *19*(4), 886-909.
464. Cox, J. C., Servátka, M., & Vadovič, R. (2017). Status quo effects in fairness games: reciprocal responses to acts of commission versus acts of omission. *Experimental Economics*, *20*(1), 1-18.
465. Belot, M., & Van de Ven, J. (2017). How private is private information? The ability to spot deception in an economic game. *Experimental Economics*, *20*(1), 19-43.
466. Kocher, M. G., Martinsson, P., Myrseth, K. O. R., & Wollbrant, C. E. (2017). Strong, bold, and kind: Self-control and cooperation in social dilemmas. *Experimental Economics*, *20*(1), 44-69.

467. Corazzini, L., Galavotti, S., Sausgruber, R., & Valbonesi, P. (2017). Allotment in first-price auctions: an experimental investigation. *Experimental Economics*, *20*(1), 70-99.
468. Gioia, F. (2017). Peer effects on risk behaviour: the importance of group identity. *Experimental Economics*, *20*(1), 100-129.
469. Alevy, J. E., & Price, M. K. (2017). Advice in the marketplace: a laboratory study. *Experimental Economics*, *20*(1), 156-180.
470. Denant-Boemont, L., Diecidue, E., & l'Haridon, O. (2017). Patience and time consistency in collective decisions. *Experimental Economics*, *20*(1), 181-208.
471. Cahlíková, J., & Cingl, L. (2017). Risk preferences under acute stress. *Experimental Economics*, *20*(1), 209-236.
472. Bosworth, S. J. (2017). The importance of higher-order beliefs to successful coordination. *Experimental Economics*, *20*(1), 237-258.
473. Bruner, D. M. (2017). Does decision error decrease with risk aversion?. *Experimental Economics*, *20*(1), 259-273.
474. Fréchette, G. R., & Yuksel, S. (2017). Infinitely repeated games in the laboratory: Four perspectives on discounting and random termination. *Experimental Economics*, *20*(2), 279-308.
475. Pan, X., & Houser, D. (2017). Social approval, competition and cooperation. *Experimental Economics*, *20*(2), 309-332.
476. Cardella, E., & Kitchens, C. (2017). The impact of award uncertainty on settlement negotiations. *Experimental Economics*, *20*(2), 333-367.
477. Blume, A., Kriss, P. H., & Weber, R. A. (2017). Pre-play communication with forgone costly messages: experimental evidence on forward induction. *Experimental Economics*, *20*(2), 368-395.
478. Chen, J., & Houser, D. (2017). Promises and lies: can observers detect deception in written messages. *Experimental economics*, *20*(2), 396-419.
479. Ryvkin, D., & Semykina, A. (2017). An experimental study of democracy breakdown, income and inequality. *Experimental Economics*, *20*(2), 420-447.
480. Baghestanian, S., Gortner, P., & Massenot, B. (2017). Compensation schemes, liquidity provision, and asset prices: an experimental analysis. *Experimental Economics*, *20*(2), 481-505.
481. Buser, T., Dreber, A., & Mollerstrom, J. (2017). The impact of stress on tournament entry. *Experimental economics*, *20*(2), 506-530.
482. Andersson, O., Miettinen, T., Hytönen, K., Johannesson, M., & Stephan, U. (2017). Subliminal influence on generosity. *Experimental Economics*, *20*(3), 531-555.
483. Sloof, R., & von Siemens, F. A. (2017). Illusion of Control and the Pursuit of Authority. *Experimental Economics*, *20*(3), 556-573.
484. Roelofs, M. R., Østbye, S. E., & Heen, E. E. (2017). Asymmetric firms, technology sharing and R&D investment. *Experimental Economics*, *20*(3), 574-600.
485. Bosch-Rosa, C., Aperjis, C., Friedman, D., & Huberman, B. A. (2017). Intolerable nuisances: some laboratory evidence on survivor curve shapes. *Experimental Economics*, *20*(3), 601-621.
486. Eriksson, T., Mao, L., & Villeval, M. C. (2017). Saving face and group identity. *Experimental Economics*, *20*(3), 622-647.
487. Jian, L., Li, Z., & Liu, T. X. (2017). Simultaneous versus sequential all-pay auctions: an experimental study. *Experimental Economics*, *20*(3), 648-669.
488. Guillen, P., & Hakimov, R. (2017). Not quite the best response: Truth-telling, strategy-proof matching, and the manipulation of others. *Experimental Economics*, *20*(3), 670-686.

489. Fenig, G., & Petersen, L. (2017). Distributing scarce jobs and output: experimental evidence on the dynamic effects of rationing. *Experimental Economics*, *20*(3), 707-735.
490. Georganas, S., Levin, D., & McGee, P. (2017). Optimistic irrationality and overbidding in private value auctions. *Experimental Economics*, *20*(4), 772-792.
491. Blanco, E., Haller, T., & Walker, J. M. (2017). Externalities in appropriation: responses to probabilistic losses. *Experimental Economics*, *20*(4), 793-808.
492. Hernandez-Lagos, P., Minor, D., & Sisak, D. (2017). Do people who care about others cooperate more? Experimental evidence from relative incentive pay. *Experimental Economics*, *20*(4), 809-835.
493. Corcos, A., Pannequin, F., & Montmarquette, C. (2017). Leaving the market or reducing the coverage? A model-based experimental analysis of the demand for insurance. *Experimental Economics*, *20*(4), 836-859.
494. Kvaløy, O., Luzuriaga, M., & Olsen, T. E. (2017). A trust game in loss domain. *Experimental Economics*, *20*(4), 860-877.
495. Koppel, L., Andersson, D., Posadzy, K., Västfjäll, D., & Tinghög, G. (2017). The effect of acute pain on risky and intertemporal choice. *Experimental economics*, *20*(4), 878-893.
496. Gee, L. K., Migueis, M., & Parsa, S. (2017). Redistributive choices and increasing income inequality: experimental evidence for income as a signal of deservingness. *Experimental Economics*, *20*(4), 894-923.
497. Hong, F., & Zhao, X. (2017). The emergence of language differences in artificial codes. *Experimental Economics*, *20*(4), 924-945.
498. Cason, T. N., Sheremeta, R. M., & Zhang, J. (2017). Asymmetric and endogenous within-group communication in competitive coordination games. *Experimental Economics*, *20*(4), 946-972.
499. Ausubel, L. M., Burkett, J. E., & Filiz-Ozbay, E. (2017). An experiment on auctions with endogenous budget constraints. *Experimental Economics*, *20*(4), 973-1006.
500. Engler, Y., Kerschbamer, R., & Page, L. (2018). Why did he do that? Using counterfactuals to study the effect of intentions in extensive form games. *Experimental Economics*, *21*(1), 1-26.
501. Evdokimov, P., & Garfagnini, U. (2018). Third-party manipulation of conflict: an experiment. *Experimental Economics*, *21*(1), 27-49.
502. Charness, G., Rustichini, A., & Van de Ven, J. (2018). Self-confidence and strategic behavior. *Experimental Economics*, *21*(1), 72-98.
503. Bosch-Rosa, C., Meissner, T., & Bosch-Domènech, A. (2018). Cognitive bubbles. *Experimental Economics*, *21*(1), 132-153.
504. Kleinlercher, D., & Stöckl, T. (2018). On the provision of incentives in finance experiments. *Experimental Economics*, *21*(1), 154-179.
505. Kesternich, M., Lange, A., & Sturm, B. (2018). On the performance of rule-based contribution schemes under endowment heterogeneity. *Experimental Economics*, *21*(1), 180-204.
506. Baillon, A., Schlesinger, H., & van de Kuilen, G. (2018). Measuring higher order ambiguity preferences. *Experimental Economics*, *21*(2), 233-256.
507. Szembrot, N. (2018). Experimental study of cursed equilibrium in a signaling game. *Experimental Economics*, *21*(2), 257-291.
508. Rubin, J., Samek, A., & Sheremeta, R. M. (2018). Loss aversion and the quantity–quality tradeoff. *Experimental Economics*, *21*(2), 292-315.
509. Bellemare, C., Sebald, A., & Suetens, S. (2018). Heterogeneous guilt sensitivities and incentive effects. *Experimental Economics*, *21*(2), 316-336.

510. Leibbrandt, A., & Lynham, J. (2018). Does the paradox of plenty exist? Experimental evidence on the curse of resource abundance. *Experimental Economics*, *21*(2), 337-354.
511. Heinrich, T., & Mayrhofer, T. (2018). Higher-order risk preferences in social settings. *Experimental economics*, *21*(2), 434-456.
512. Cooper, D. J., & Van Huyck, J. (2018). Coordination and transfer. *Experimental Economics*, *21*(3), 487-512.
513. Van Huyck, J., & Stahl, D. O. (2018). Conditional behavior and learning in similar stag hunt games. *Experimental Economics*, *21*(3), 513-526.
514. Van Huyck, J., Viriyavipart, A., & Brown, A. L. (2018). When less information is good enough: experiments with global stag hunt games. *Experimental Economics*, *21*(3), 527-548.
515. Leng, A., Friesen, L., Kalayci, K., & Man, P. (2018). A minimum effort coordination game experiment in continuous time. *Experimental Economics*, *21*(3), 549-572.
516. Duffy, J., & Fehr, D. (2018). Equilibrium selection in similar repeated games: Experimental evidence on the role of precedents. *Experimental Economics*, *21*(3), 573-600.
517. Kurz, V., Orland, A., & Posadzy, K. (2018). Fairness versus efficiency: how procedural fairness concerns affect coordination. *Experimental Economics*, *21*(3), 601-626.
518. Jacquemet, N., Luchini, S., Shogren, J. F., & Zylbersztejn, A. (2018). Coordination with communication under oath. *Experimental Economics*, *21*(3), 627-649.
519. Filiz-Ozbay, E., Ham, J. C., Kagel, J. H., & Ozbay, E. Y. (2018). The role of cognitive ability and personality traits for men and women in gift exchange outcomes. *Experimental Economics*, *21*(3), 650-672.
520. Camera, G., & Casari, M. (2018). Monitoring institutions in indefinitely repeated games. *Experimental Economics*, *21*(3), 673-691.
521. Arifovic, J., & Ledyard, J. (2018). Learning to alternate. *Experimental Economics*, *21*(3), 692-721.
522. Zhou, W., & Hey, J. (2018). Context matters. *Experimental economics*, *21*(4), 723-756.
523. Regner, T. (2018). Reciprocity under moral wiggle room: Is it a preference or a constraint?. *Experimental Economics*, *21*(4), 779-792.
524. He, T. S., & Hong, F. (2018). Risk breeds risk aversion. *Experimental Economics*, *21*(4), 815-835.
525. Chen, J. I., & Kamei, K. (2018). Disapproval aversion or inflated inequity acceptance? The impact of expressing emotions in ultimatum bargaining. *Experimental economics*, *21*(4), 836-857.
526. Wu, J. (2018). Indirect higher order beliefs and cooperation. *Experimental Economics*, *21*(4), 858-876.
527. Clingingsmith, D., & Sheremeta, R. M. (2018). Status and the demand for visible goods: Experimental evidence on conspicuous consumption. *Experimental Economics*, *21*(4), 877-904.
528. Aycinena, D., & Rentschler, L. (2018). Auctions with endogenous participation and an uncertain number of bidders: experimental evidence. *Experimental Economics*, *21*(4), 924-949.

Journal of the European Economic Association

529. Goeree, J. K., Plott, C. R., & Wooders, J. (2004). Bidders' choice auctions: Raising revenues through the right to choose. *Journal of the European Economic Association*, *2*(2-3), 504-515.
530. Gächter, S., & Thöni, C. (2005). Social learning and voluntary cooperation among like-minded people. *Journal of the European Economic Association*, *3*(2-3), 303-314.
531. Bohnet, I., Harmgart, H., & Tyran, J. R. (2005). Learning trust. *Journal of the European Economic Association*, *3*(2-3), 322-329.
532. Brandts, J., & Cooper, D. J. (2007). It's what you say, not what you pay: An experimental study of manager-employee relationships in overcoming coordination failure. *Journal of the European Economic Association*, *5*(6), 1223-1268.
533. Fehr, E., Zehnder, C., & Hart, O. (2009). Contracts, reference points, and competition—behavioral effects of the fundamental transformation. *Journal of the European Economic Association*, *7*(2-3), 561-572.
534. Cooper, D. J., & Kagel, J. H. (2009). The role of context and team play in cross-game learning. *Journal of the European Economic Association*, *7*(5), 1101-1139.
535. Goeree, J. K., Palmer, K., Holt, C. A., Shobe, W., & Burtraw, D. (2010). An experimental study of auctions versus grandfathering to assign pollution permits. *Journal of the European Economic Association*, *8*(2-3), 514-525.
536. Müller, W., & Schotter, A. (2010). Workaholics and dropouts in organizations. *Journal of the European Economic Association*, *8*(4), 717-743.
537. Abeler, J., Altmann, S., Kube, S., & Wibral, M. (2010). Gift exchange and workers' fairness concerns: when equality is unfair. *Journal of the European Economic Association*, *8*(6), 1299-1324.
538. Ericson, K. M. M. (2011). Forgetting we forget: Overconfidence and memory. *Journal of the European Economic Association*, *9*(1), 43-60.
539. Grosse, S., Putterman, L., & Rockenbach, B. (2011). Monitoring in teams: using laboratory experiments to study a theory of the firm. *Journal of the European Economic Association*, *9*(4), 785-816.
540. Carpenter, J. P., & Matthews, P. H. (2012). Norm enforcement: anger, indignation, or reciprocity?. *Journal of the European Economic Association*, *10*(3), 555-572.
541. Hyndman, K., Ozbay, E. Y., Schotter, A., & Ehrblatt, W. Z. E. (2012). Convergence: an experimental study of teaching and learning in repeated games. *Journal of the European Economic Association*, *10*(3), 573-604.
542. Brown, M., Falk, A., & Fehr, E. (2012). Competition and relational contracts: the role of unemployment as a disciplinary device. *Journal of the European Economic Association*, *10*(4), 887-907.
543. Shurchkov, O. (2012). Under pressure: gender differences in output quality and quantity under competition and time constraints. *Journal of the European Economic Association*, *10*(5), 1189-1213.
544. Krupka, E. L., & Weber, R. A. (2013). Identifying social norms using coordination games: Why does dictator game sharing vary?. *Journal of the European Economic Association*, *11*(3), 495-524.
545. Fischbacher, U., & Föllmi-Heusi, F. (2013). Lies in disguise—an experimental study on cheating. *Journal of the European Economic Association*, *11*(3), 525-547.
546. Gächter, S., Nosenzo, D., & Sefton, M. (2013). Peer effects in pro-social behavior: Social norms or social preferences?. *Journal of the European Economic Association*, *11*(3), 548-573.

547. Cappelen, A. W., Moene, K. O., Sørensen, E. Ø., & Tungodden, B. (2013). Needs versus entitlements—an international fairness experiment. *Journal of the European Economic Association*, *11*(3), 574-598.
548. Bartling, B., Fehr, E., & Schmidt, K. M. (2013). Use and abuse of authority: a behavioural foundation of the employment relation. *Journal of the European Economic Association*, *11*(4), 711-742.
549. Serra-Garcia, M., Van Damme, E., & Potters, J. (2013). Lying about what you know or about what you do?. *Journal of the European Economic Association*, *11*(5), 1204-1229.
550. Choi, S., & Lee, J. (2014). Communication, coordination, and networks. *Journal of the European Economic Association*, *12*(1), 223-247.
551. Durante, R., Putterman, L., & Van der Weele, J. (2014). Preferences for redistribution and perception of fairness: An experimental study. *Journal of the European Economic Association*, *12*(4), 1059-1086.
552. Baeriswyl, R., & Cornand, C. (2014). Reducing overreaction to central banks' disclosures: theory and experiment. *Journal of the European Economic Association*, *12*(4), 1087-1126.
553. Heffetz, O., & List, J. A. (2014). Is the endowment effect an expectations effect?. *Journal of the European Economic Association*, *12*(5), 1396-1422.
554. Fehr, E., Hart, O., & Zehnder, C. (2015). How do informal agreements and revision shape contractual reference points?. *Journal of the European Economic Association*, *13*(1), 1-28.
555. Bartling, B., & Schmidt, K. M. (2015). Reference points, social norms, and fairness in contract renegotiations. *Journal of the European Economic Association*, *13*(1), 98-129.
556. Benoît, J. P., Dubra, J., & Moore, D. A. (2015). Does the better-than-average effect show that people are overconfident?: Two experiments. *Journal of the European Economic Association*, *13*(2), 293-329.
557. Ewers, M., & Zimmermann, F. (2015). Image and misreporting. *Journal of the European Economic Association*, *13*(2), 363-380.
558. Kimbrough, E. O., & Vostroknutov, A. (2016). Norms make preferences social. *Journal of the European Economic Association*, *14*(3), 608-638.
559. Brandts, J., Ellman, M., & Charness, G. (2015). Let's talk: How communication affects contract design. *Journal of the European Economic Association*, *14*(4), 943-974.
560. Grossman, Z., & Van der Weele, J. J. (2017). Self-image and willful ignorance in social decisions. *Journal of the European Economic Association*, *15*(1), 173-217.
561. Cappelen, A. W., Halvorsen, T., Sørensen, E. Ø., & Tungodden, B. (2017). Face-saving or fair-minded: What motivates moral behavior?. *Journal of the European Economic Association*, *15*(3), 540-557.
562. Dertwinkel-Kalt, M., Köhler, K., Lange, M. R., & Wenzel, T. (2017). Demand shifts due to salience effects: Experimental evidence. *Journal of the European Economic Association*, *15*(3), 626-653.
563. Choi, S., Galeotti, A., & Goyal, S. (2017). Trading in networks: theory and experiments. *Journal of the European Economic Association*, *15*(4), 784-817.
564. Gneezy, U., Goette, L., Sprenger, C., & Zimmermann, F. (2017). The limits of expectations-based reference dependence. *Journal of the European Economic Association*, *15*(4), 861-876.
565. Barrett, S., & Dannenberg, A. (2017). Tipping versus cooperating to supply a public good. *Journal of the European Economic Association*, *15*(4), 910-941.

566. d'Adda, G., Darai, D., Pavanini, N., & Weber, R. A. (2017). Do leaders affect ethical conduct?. *Journal of the European Economic Association*, *15*(6), 1177-1213.
567. Aghion, P., Fehr, E., Holden, R., & Wilkening, T. (2018). The Role of Bounded Rationality and Imperfect Information in Subgame Perfect Implementation—An Empirical Investigation. *Journal of the European Economic Association*, *16*(1), 232-274.
568. Herz, H., & Taubinsky, D. (2018). What makes a price fair? An experimental study of transaction experience and endogenous fairness views. *Journal of the European Economic Association*, *16*(2), 316-352.
569. Agranov, M., Goeree, J. K., Romero, J., & Yariv, L. (2018). What makes voters turn out: The effects of polls and beliefs. *Journal of the European Economic Association*, *16*(3), 825-856.

3. Is your privacy for sale? An experiment on the willingness to reveal sensitive information

This chapter has been published as:

Cloos, J., Frank, B., Kampenhuber, L., Karam, S., Luong, N., Möller, D., Monge-Larrain, M., Tan Dat, N., Nilgen, M., & Rössler, C. (2019). Is Your Privacy for Sale? An Experiment on the Willingness to Reveal Sensitive Information. *Games*, *10*(3), 28.

This paper can be downloaded via

https://doi.org/10.3390/g10030028

Abstract: We investigate whether individuals' self-stated privacy behavior is correlated with their reservation price for the disclosure of personal and potentially sensitive information. Our incentivized experiment has a unique setting: Information about choices with real implications could be immediately disclosed to an audience of fellow first semester students. Although we find a positive correlation between respondents' willingness to accept (WTA) disclosure of their private information and their stated privacy behavior for some models, this correlation disappears when we change the specification of the privacy index. Independent of the privacy index chosen we find that the WTA is significantly influenced by individual responses to personal questions, as well as by different decisions to donate actual money, indicating that the willingness to protect private information depends on the delicacy of the information at stake.

Keywords: privacy, personal data, social disapproval, WTA experiment, donation experiment

3.1 Introduction

Privacy is a fundamental human right; but is yours for sale? This is a question that consumers, especially of seemingly 'free' online services, face more and more frequently. Oftentimes, people state that their privacy is highly important to them, but when revealing their actual behavior, e.g., when using social networks or mobile tracking devices, private information is disclosed without further ado. This intention-action gap between stated preferences and actual behavior, commonly referred to as 'privacy paradox', has intrigued researchers from different scientific fields since the early 2000s (see Acquisti et al. (2015) for an overview). Dealing with the issue of privacy is generally difficult: Privacy is subjective, idiosyncratic and intertemporal; its value varies between different people and different information, and it may change over time (Acquisti et al., 2016). This heterogeneity might be one reason why empirical evidence on the existence of the privacy paradox so far is mixed: Some of the literature supports its existence (e.g., Taddicken, 2014; Acquisti and Gross, 2006; Acquisti and Grossklags, 2005; Spiekermann et al., 2001), while others provide evidence challenging it (e.g., Blank et al., 2014; Egelman et al., 2013; Son and Kim, 2008).

Economic experiments attempting to measure the monetary valuation of privacy do not aim to examine the privacy paradox problem directly. In the current literature, there is no standardized method of measuring the monetary value of privacy, and, hence, the degree of heterogeneity in findings is quite substantial. As pointed out in Benndorf and Normann (2018) and Tsai et al. (2011), the mixed evidence on privacy valuation can be explained by three fundamental issues within the study setting: Incentives, saliency, and transparency. First, not all experimental studies were conducted in incentivized settings (John et al., 2011). Secondly, some studies emphasized the topic of privacy throughout the experiment, while others made privacy issues less salient. Finally, whereas the uses of the data collected in some studies were clearly mentioned and transparent (Benndorf and Normann, 2018; Plesch and Wolff, 2018), in other studies, this information was kept ambiguous.

Another explanation for the diverse results in the privacy literature is the nature of the private information potentially to be revealed. Some studies focus on the reservation price (or respondents' willingness to accept - WTA) disclosure of individual sensitive information, e.g., contact data (Plesch and Wolff, 2018), private address information, or *Facebook* data (Benndorf and Normann, 2018), whereas other studies focus on the willingness to reveal one's

bad deeds (John et al., 2011).[18] In the experiment presented in this paper, we elicit both aforementioned types of information and additionally investigate whether people discriminate between their valuations for keeping "good" and "bad" decisions private. Moreover, while some studies correlate the desirability of a trait and the WTA for the information disclosure (see Huberman et al., 2005), to our knowledge there are as yet no monetary incentivized studies investigating information disclosure regarding morally afflicted actual behavior (e.g., donating or taking to/from an organization). When asking for reservation prices for the willingness to disclose data, from e.g., social networks or online shopping decisions, it is unclear how the actual content of the data potentially to be revealed differs between individuals. For example, it seems reasonable to assume that people on average would ask a higher reservation price for the disclosure of their data if there are many photos from a beach party on their *Facebook* timeline than if there are just a couple of photos from a hiking trip. Studies in which *Facebook* data or online shopping decisions can be revealed undoubtedly have a high degree of relevance to reality. However, it is difficult to control for content differences when analyzing such data. Therefore, this paper contributes and expands the existing literature by using an experimental design that monetarily quantifies the value that is individually assigned to the information that can potentially be revealed.

In the present study, we examine whether the privacy paradox holds when individuals decide about revealing potentially sensitive information about actual behavior. In addition to that, our study considers the incentive, saliency and transparency issues raised by Benndorf and Normann (2018). We conducted an experiment with first semester students during their university introduction week. First, we collected information regarding the subjects' online privacy behavior, as well as their general attitude towards privacy concerns. Afterwards, we collected some private information from the participants, before they were given the opportunity to donate or take from three different organizations (henceforth DOT decisions). By including different types of organizations, we made sure that the DOT decisions had varying degrees of moral loading, potentially inducing image concerns in case they would be publicly revealed in front of the audience of fellow first semester students. In the final step of the experiment, we employed the Becker-DeGroot-Marschak mechanism (henceforth BDM; Becker et al., 1964) to elicit both the subjects' minimum WTA to have their private information

[18] In John et al. (2011), subjects fill out surveys with questions about their sexual behavior, among other things, and can provide their email address if they want to receive an evaluation. The study finds that a different survey header design influences the contents of the answers and the willingness to provide one's own email address. These results are consistent with a study by Chavanne (2018), whose findings also suggest that the willingness to disclose personal data is context-specific.

revealed (henceforth $WTA_{details}$), as well as their minimum WTA to have their DOT decisions revealed (henceforth WTA_{DOT}).

We find that self-stated privacy behavior is positively and significantly correlated with both $WTA_{details}$ and WTA_{DOT} when we use an additive index for privacy. We no longer observe a significant correlation when we calculate privacy indices based on factor analysis. Hence, we are not able to find conclusive evidence for the privacy paradox. We additionally detect that there is no significant difference between the $WTA_{details}$ and WTA_{DOT}. As expected, we find significant differences between DOT decisions for different organizations.

The remainder of this paper is organized as follows. Section 3.2 reviews relevant literature; section 3.3 describes the experimental design, hypotheses and procedures; section 3.4 presents the results and a robustness check; and section 3.5 concludes.

3.2 Related literature

Today, privacy is a more important topic than ever. Overall advances in information technology and the widespread distribution of the internet allows companies to collect vast amounts of user data, which in turn are combined and analyzed to derive valuable insights about markets, as well as individual consumers preferences. In this global digital environment, where personal data is increasingly used as kind of currency to pay for otherwise 'free' online services, more and more questions are being raised regarding the monetary valuation of personal data and privacy. As Acquisti et al. (2016) point out, both the protection and sharing of personal information induces costs and benefits for individuals, as well as for society as a whole.

Our paper adds to the literature of the willingness to pay for privacy. Most closely related to our setting is the experiment by Schudy and Utikal (2017). They used the BDM to evaluate the willingness to share individuals' private information with a varying share of other students. By collecting information about the participants' address and body dimensions, they found that an increasing number of potential information receivers decreases the willingness to share one's personal data. The social distance between the subjects only mattered for women with regards to their body dimensions. Whether the information to be disclosed is verified only had an effect on men. Another closely related study is Benndorf and Normann (2018), who used the BDM to evaluate the willingness to share individuals' personal data with a telecommunications company. They varied the potential information which could be transmitted to the telecommunications company and find that individuals have a higher willingness to sell contact data, compared to data from their *Facebook* account. Another finding is that a share of 10 to

20% of subjects refuse in general to sell their data, whereas a comparable share is willing to sell its data for a very small amount.

Various studies tried to investigate peoples' monetary valuation of keeping personal information private. In one of the earliest empirical contributions to this issue, Huberman et al. (2005) conducted a series of experimental auctions. They used information about their participants' weight and age and asked them about the minimum amount they would be willing to accept in order to make that information public. Their findings reveal that the value of privacy depends on the type of information and that there is a higher valuation of the weight information if perceived as embarrassing. While Huberman et al. (2005) used auctions, other studies employed different techniques to elicit the valuation of privacy. These include choice menus (Dogruel et al., 2017; Egelman et al., 2013; Beresford et al., 2012), the already mentioned BDM (Benndorf and Normann, 2018; Schudy and Utikal, 2017), open questions Fuller (2019) or the indication of a fixed amount of money (Benndorf and Normann, 2018; Plesch and Wolff, 2018). While in some empirical studies (Benndorf and Normann, 2018; Schudy and Utikal, 2017) personal data of participants could be disclosed, in other studies (Benndorf 2018; Schudy and Utikal, 2018) the data to be disclosed were endogenously assigned to the participants in the experiment. Furthermore, various studies in the field of mechanism design (Jin et al., 2018; Niu et al., 2018; Ghosh and Roth, 2015) address the issue of adequate privacy compensation on the internet.

Other experimental studies tried to tackle the 'privacy paradox' directly. Spiekermann et al. (2001) compared self-reported privacy preferences with actual behavior during online shopping and found that people tend to reveal a discrepancy between their stated privacy preferences and actual behavior, thereby providing evidence in favor of the 'privacy paradox'. This is in line with the findings of Beresford et al. (2012), who conducted a field experiment in which individuals could either purchase DVDs from an online store where they had to reveal their monthly income and their date of birth, or from an online store where they had to reveal their favorite color and their year of birth. The shops differed in that DVDs in the first online store were one euro cheaper. They found that significantly more individuals purchased the cheaper DVDs, although after the experiment, they stated high privacy concerns, thereby revealing behavior in line with the 'privacy paradox'. Contrasting evidence is provided by Dogruel et al. (2017), who tested the willingness to pay additional money for smartphone apps if the app did not reveal private information to third parties. They find that if the default app exhibits a high level of privacy, individuals tend to buy a high-level privacy setting with a

higher likelihood. In addition, they find that individuals who stated a high concern for privacy before the experiment, are more likely to purchase the app with a high privacy level. Kokolakis (2017) reviews the literature on studies investigating the privacy paradox and explains the diversity in research results with the heterogeneity in research methods and contexts, as well as different conceptualizations of the 'privacy paradox', while generally arguing for revealed - instead of stated - preference methods.

Potential explanations for exhibiting behavior in line with the 'privacy paradox' are consumer convenience (e.g., a smartphone application that works better by sharing the location information), preferences for saving time (e.g., users who do not spend sufficient time looking for safer options) or to have better access to social networks (e.g., *Facebook*) (Hann et al., 2007).

3.3 Method

3.3.1 Experimental design

The experiment consisted of six stages. In the instructions (see Appendix, section 3.6.1), we informed participants about the rules of the experiment and that three participants would be randomly selected to receive payments of at least €100 each at the end of the experiment. We also asked the participants to answer all questions truthfully and according to their personal opinion. In the second stage, participants answered a short questionnaire about their attitudes regarding privacy and their personal behavior related to privacy and data security on the internet. The questionnaire contained seven items which had to be answered on a five-point Likert scale. As far as we know, no adequate index for online privacy behavior exists yet. We, therefore, created a provisional index with only a few items on real life privacy behavior that should theoretically cover the concept. In order to avoid measurement artifacts through acquisition, we rescaled three of the items so that the subjects achieve a lower index value through more agreement.

During the third stage, we asked the participants three questions on 'personal details' which were not related to each other. We tried to elicit information which could potentially be uncomfortable to reveal in the presence of others. At the same time, we did not want to make the content of the answers so unpleasant as to reveal that subjects would possibly not sell at any price. The questions were: (1) 'Would you like to lose weight?' (henceforth, weight question); (2) 'Do you smoke?' (henceforth, smoke question); and (3) 'What was the average final grade of your last school leaving certificate (e.g., Abitur)?' (henceforth, grade question). In this stage,

as well as in all other relevant stages of the experiment, we made it clear that those participants with invalid or missing answers would lose their chance of getting paid. In stage 2 (privacy questionnaire), as well as in stage 3 (personal details) we tried to select the questions in a way that they seemed to be somewhat familiar to a majority of our participants. At the same time, we only asked a limited number of questions in order to keep the experiment simple and concise and to ensure the attention of our participants.

The fourth stage contained the DOT decisions. The purpose of this stage was to collect data on monetary allocation decisions with a moral component. First, we informed the participants that they were endowed with an amount of €150 and that three organizations were endowed with amounts of €50 each. Afterwards, the subjects could choose to donate to or to take from each organization any amount between €0 and €50, whereas only one decision was to be realized in the end. Participants could also explicitly indicate that they neither wanted to donate nor take at all. We used the strategy method (Selten, 1967) and asked the participants to make allocation decisions when matched to three different organizations: The KSV Hessen Kassel (Kasseler Sport-Verein; a local soccer club), the ILGA (International Lesbian, Gay, Bisexual, Trans and Intersex Association), and the UNHCR (United Nations Refugee Agency). Undoubtedly, endowing all individuals with an amount of €150 might affect their DOT decisions. However, since all participants were endowed with the same amount, this would only result in a level effect, whereas differences in the DOT decisions can be attributed to participants' individual characteristics.

As the results of List (2007) show, donations in a dictator game are predominantly zero as soon as the subjects have a "take" option in addition to a "donate" option. Hence, we used different organizations with different goals in order to increase the chance that there would be a variation in the DOT decisions between the participants. Benndorf and Normann (2018) point out that it is unclear how and if students will handle the information made public about other students. Since in our experiment, the participants were first semester students, we assumed that there would be particularly strong image concerns involved, because of the importance of making a good first impression.

In the fifth stage, the participants were introduced to the possibility of selling information about their previous statements and decisions (stage three and four). We, therefore, used the BDM, which is a standard incentive compatible method for eliciting participants' WTA. We offered subjects the possibility to buy the right to read out publicly two pieces of information: The minimum amount that we should pay them to reveal (a) their answers in the "personal details"

section ($WTA_{details}$), and (b) their DOT decisions (WTA_{DOT}) in front of the audience. We determined our maximum offer for each piece of information by randomly choosing one of three envelopes containing possible prices. If our maximum offer for a piece of information exceeded the valuation stated by the participant, then she (or he) would be paid the amount that was written on the sheet inside the envelope, and her information would be read out publicly, while she had to stand up to guarantee that all participants could see her. To guarantee that the participants understood the BDM mechanism, we demonstrated it beforehand by illustrating an exemplary transaction. In the instructions, we clearly stated that entering WTAs above €100 would result in the non-disclosure of the information in question. Thus, similar to Schudy and Utikal (2017) and Benndorf and Normann (2018), the random draw of the BDM in our experiment had an upper bound. Bohm et al. (1997) find that specifying an upper bound when using the BDM can have an anchoring effect on the decisions of the subjects. However, as Benndorf and Normann (2018) point out, this anchoring can work in both directions. Thus, we decided to specify an upper bound for WTAs, since otherwise, it might be difficult for subjects to state reasonable sales prices for the requested pieces of information. As Ghosh and Roth (2015) point out, the indication of a reservation price for the disclosure of private information already constitutes a disclosure of private information. Applied to our experiment, this implies that participants could indicate a higher WTA in order not to appear as if their privacy is not of great value to them.

The final stage consisted of a brief questionnaire for getting feedback from the experiment and collecting basic demographic information.

3.3.2 Hypotheses

We used our experimental design to test three main hypotheses. Our first hypothesis follows the findings of Kokolakis (2017), that there is mixed evidence about the existence of a privacy paradox. Thus, we test whether the paradox holds in our setting, which would be reflected by a higher index of self-stated privacy behavior not affecting the $WTA_{details}$ and WTA_{DOT} decisions. Thus, our first main hypothesis is:

Hypothesis 1 (H1). The index of self-stated privacy behavior is not positively correlated with the $WTA_{details}$ and WTA_{DOT}.

For the next hypothesis, we focus on the answers in the ‘personal details’ section. Regarding the weight-question, we expect a positive answer to lead to an increase in $WTA_{details}$. The possibility that the stated desire to lose weight will be read out publicly is likely to create an

unpleasant feeling among many participants. Regarding the smoke-question, again, we expect a positive response to lead to an increase in $WTA_{details}$, since there is a social norm against smoking. For the grade-question, we also expect that the indication of a poor final school grade will lead to an increase in $WTA_{details}$. Therefore, our next hypothesis is:

Hypothesis 2 (H2). Stating the desire to lose weight, being a smoker and indicating a poor high school grade is positively correlated with $WTA_{details}$.

Our final hypothesis is related to the participants' DOT decisions. Participants were already endowed with a sizeable monetary amount at the start of the DOT decision section. This would have made a take-decision at least appear greedy or even brash, leading to an incentive to hide such decisions. Thus, we expect that taking will lead to an increase in WTA_{DOT}. Conversely, giving an additional amount to one of the organizations would have most likely been regarded as generous or even as exemplary, so people showing such behavior would have had less incentive to hide such decisions. Thus, we assume that donating to the organizations would generally induce a decrease in WTA_{DOT}. The corresponding hypothesis is:

Hypothesis 3 (H3). Taking (giving) from (to) the individual organizations, leads to an increase (decrease) in WTA_{DOT}.

3.3.3 Experimental procedure

The experiment took place at the University of Kassel during the introductory week for new students (October 2018). We ran two sessions of approximately 30 min with pen and paper, and in each session, only three randomly chosen participants received a real payoff. In total 105 students (44.76% female, 54.29% male, 0.95% not specified) participated in the experiment. In the first session, we had 61 freshmen from the Economics and Engineering Bachelor Program and in the second session 44 from Politics, Sociology or History with a minor in economics. The average age of the participants was 20.7 years, whereas 71.84% were younger than 22. Due to incomplete questionnaires, we had to drop six observations. Subsequently, in total, we have 99 observations for our analysis. All participants got a small gift (bike saddle cover or a pen) for participating in the study. The average payoff per person was €8.94. Based on the results of the experiment, we donated €0.00 to the KSV, €165.00 to the ILGA and €200.00 to the UNHCR. As stated in the instructions, we published the donated amounts on the university

homepage of Björn Frank.[19] In order to prevent participants from inferring the decisions of the winners, we added a small additional amount of money to the individual donations.

To preserve anonymity, we gave each participant a card with an identification number and instructions about when and where to collect their respective payment. We distributed sheets of paper with instructions for the first three stages. Once everybody had filled out the respective questions of these stages, we collected the completed forms. After that, we explained and demonstrated the BDM. We then distributed instructions for the remaining stages. After collecting the completed questionnaires, one of the students randomly drew the two maximum offers for the right to reveal personal details and the DOT decision. Afterwards, the three participants who received the actual payment and the organizations that would match each of the selected participants were randomly drawn. We only revealed the information of the participants whose reservation prices were lower than the selected prices. In session one, we revealed the personal details of one winner. In session two, we revealed the personal details, as well as the DOT decisions of one winner.

3.4 Results

3.4.1 Descriptive statistics

Average responses for questions indicating individuals' privacy preferences are reported in Table 3.1. We distinguish between self-stated privacy behavior (questions a1–a5) and stated privacy attitudes (questions a6–a7). Since we find low variation for privacy attitudes, we constructed an index for privacy concerns consisting of an average of all five privacy behavior questions (a1–a5).

[19] See https://www.uni-kassel.de/fb07/fileadmin/datas/fb07/5-Institute/IVWL/Frank/teaching/diverse_bf/Spendennachweis.pdf (accessed June 27, 2019).

Question	Mean (sd)
a1. Some websites use special tools to identify returning users and to provide them with personalized information (such as advertisements). If a website offered you such a service, would you agree? [1,4]	3.53 (1.17)
a2. How often do you use the private mode (incognito mode) of your web browser? [2]	2.42 (1.15)
a3. Do you use your real name on social networks (like *Facebook*)? [2,4]	2.13 (1.23)
a4. Do you allow apps on your smartphone, tablet or laptop to determine your location and record it if necessary? [2,4]	3.03 (1.09)
a5. Do you cover up your laptop's webcam? [2]	2.83 (1.76)
a6. How important is privacy to you in general? [3]	3.69 (0.72)
a7. How important is the security of your personal data to you? [3]	4.01 (0.96)

Table 3.1: Average scores for privacy preference indicators.

Notes. [1] We used a five-point Likert scale where 1 means "in no case", 2 "rather not", 3 "I don't know", 4 "rather yes", 5 "in any case". [2] We used a five-point Likert scale where 1 means "never", 2 "rarely", 3 "sometimes", 4 "often", 5 "always". [3] We used a five-point Likert scale where 1 means "unimportant", 2 "rather unimportant", 3 "important", 4 "rather important", 5 "very important". [4] Since our goal was to indicate preferences for privacy, we recoded the outcomes of these questions in the opposite direction to have all questions on a common scale.

The mean amount which was donated/taken was €−3.64 to KSV, €16.62 to ILGA and €34.80 to UNHCR. All means are significantly different from each other (all p-values < 0.01) and we, therefore, see strong signs of discrimination between the three organizations by our participants. As can be seen in Figure 3.1, the distribution of amounts donated/taken varied strongly between the three organizations. Among our participants a share of 38% decided to take, and 32% decided to neither take nor donate to the KSV; 15% decided to take, and 22% decided to neither take nor donate to the ILGA; and 5% decided to take, and 10% decided to neither take nor donate to the UNHCR.

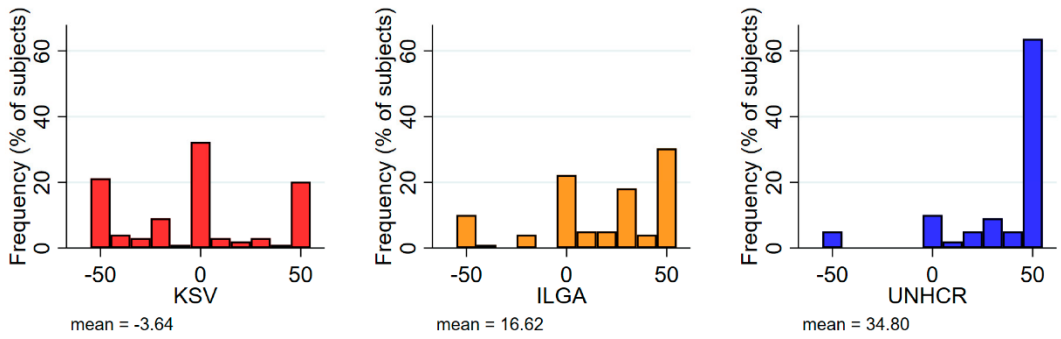


Figure 3.1: Distribution of DOT decisions for each organization.

Over all the organizations, a share of 45.45% (88.89%) of our participants decided to take (donate) from (to) at least one organization, and 10.10% (66.67%) of our participants decided to take (donate) money from (to) at least two organizations. The mean amounts for WTA_{DOT} were never significantly different within these groups (all p-values > 0.1). In section 3.4.2, we

conduct a more detailed analysis of amounts donated/taken to/from the individual organizations.

Figure 3.2 shows distributions for $WTA_{details}$ and WTA_{DOT}. The average $WTA_{details}$ was €48.17, and the average WTA_{DOT} was €50.45. Eleven individuals (10.89%) chose a $WTA_{details}$ that clearly indicates their preference to exclude the possibility that their personal details would be revealed, whereas 13 (12.87%) indicated the same for their WTA_{DOT}. It was also clear to participants that any stated value above €100 would definitely not lead to a disclosure of their personal details or their DOT decisions in the experiment. Since we did not want to lose these observations and we did not want outliers to bias our results substantially, we substituted all prizes above €100 with €101 (in the following we also conducted tobit regressions). A Wilcoxon matched-pairs signed-ranks test revealed that there are no significant differences between the distribution of $WTA_{details}$ and WTA_{DOT} ($p = 0.300$).[20]

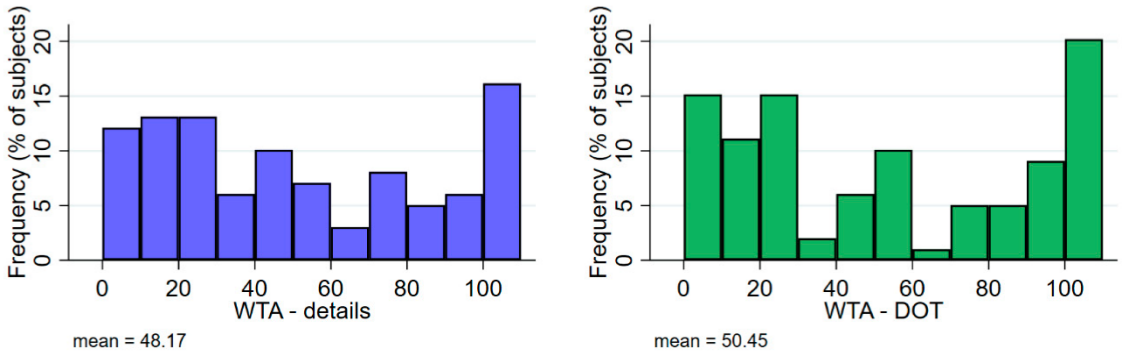


Figure 3.2: Distribution of willingness to accept (WTA) bids across personal details and DOT decisions.

3.4.2 Main regression results

Table 3.2 shows the results for an OLS-regression using robust standard errors where the dependent variable is either the $WTA_{details}$ (model 1–3) or the WTA_{DOT} (model 4–6). The privacy variable in Table 3.2 is constructed from the additive index of all privacy behavior questions (a1–a5) described in Section 3.4.1. The results of these OLS regressions do not support H1, since we find that an increase of the index for privacy behavior by one unit on average results in an increase of $WTA_{details}$ between ~€11.33 and ~€12.07, and an increase of WTA_{DOT} between ~€9.73 and ~€9.78. This effect is robust across different model specifications, although the significance and effect size drop slightly. However, we observe this effect is no longer significant in model (6). Additionally, we find only partial support for H2. Stating the desire to lose weight increases the corresponding $WTA_{details}$ roughly between

[20] If we only look at the WTAs below €101, the average values are even closer to each other ($WTA_{details}$ = €41.84; WTA_{DOT} = €42.43; $p = 0.380$).

~€13.60 and ~€16.20. However, we find no support for our hypothesis that any information besides the desire to lose weight influences $WTA_{details}$. Stating that one smokes, or a poor high school grade does not significantly influence $WTA_{details}$. As already pointed out in the descriptive statistic section, we find a significantly differing donation behavior among the different organizations. The overall donation is the highest for the UNHCR, followed by ILGA and then KSV. The overall donation-decisions are significantly different from each other (Wilcoxon matched-pair test, $p < 0.01$). Furthermore, for model (5) in Table 3.2 we observe that decreasing the donated amount to the ILGA by one euro increases the minimum WTA_{DOT} by €0.26. So, we find evidence for H3 only concerning transfers to/from ILGA, but not for the other two organizations.

	(1) $WTA_{details}$	(2) $WTA_{details}$	(3) $WTA_{details}$	(4) WTA_{DOT}	(5) WTA_{DOT}	(6) WTA_{DOT}
privacy	12.07**	11.33**	11.56**	9.73*	9.78*	7.99
	(4.98)	(4.89)	(4.95)	(5.84)	(5.68)	(5.87)
weight		13.63*	16.17**			
		(7.46)	(7.89)			
smoke		7.24	6.57			
		(8.81)	(8.93)			
grade		−2.13	−1.53			
		(6.06)	(6.07)			
KSV					0.06	0.07
					(0.11)	(0.11)
ILGA					−0.26**	−0.23
					(0.13)	(0.14)
UNHCR					0.03	0.03
					(0.17)	(0.17)
age			−0.47			2.59*
			(1.53)			(1.41)
female			−4.03			6.11
			(7.74)			(7.92)
session1	8.08	9.61	8.66	13.71*	11.13	11.98
	(7.21)	(7.61)	(8.19)	(7.64)	(7.94)	(8.54)
Constant	9.91	9.71	19.08	14.41	19.49	−32.87
	(15.46)	(24.03)	(34.86)	(17.56)	(18.53)	(30.67)
Observations	94	94	92	94	94	92
F	3.55	2.41	1.88	2.84	2.22	2.70
R^2	0.08	0.12	0.13	0.07	0.11	0.13

Table 3.2: OLS regressions of WTA on self-stated privacy behavior (additive index of a1–a5) and information at stake.

Notes. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Out of 99 individuals, we lose five in the first four model specifications since some participants chose "I don't know/don't want to answer" for some of the privacy questions. Additionally, we lose two more observations in model (5) and (6) because two participants did not report their age. KSV, Kasseler Sport-Verein Hessen Kassel, (a local soccer club); ILGA, International Lesbian, Gay, Bisexual, Trans and Intersex Association; UNHCR, United Nations Refugee Agency.

3.4.3 Robustness

Since our dependent variable for our main regression had lower (0) and upper (100) bounds, we also checked whether our results hold in a censored tobit regression, where we find that

effect sizes and significances increase slightly (Appendix Table 3.3). However, instead of using an additive index for privacy concerns, we also conducted factor analysis and tested our hypotheses again using indices which contained different factor loadings of all questions regarding privacy (a1–a7, Cronbach's alpha = 0.54; Appendix Table 3.4) and also double checked by only using an index that provided us with the highest Cronbach's alpha (a1, a3, a6, a7; alpha = 0.64; Appendix Table 3.5) to see whether our results hold across different specifications for privacy concerns. Although both indices constructed with factor analysis do not meet the critical value to be considered acceptable measures for a common underlying concept, we observe that all four variable specifications correlate reasonably well with each other. Thus, we are quite confident that we actually measure the same underlying concept (Appendix Table 3.6). We observe that our results change substantially if we use different variable specifications for privacy concerns since we no longer observe significant correlations between our indices and WTA's, and, therefore, our results no longer falsify the existence of a privacy paradox in our setting (Appendix Tables 3.7 and 3.8). Additionally, we find that the largest part of our result from the first regression using the additive index was driven by question a5, which asked participants whether they cover their webcam (Appendix Table 3.9). This result is not surprising, as the question of webcam coverage is more likely than the other questions to be answered with a very high or very low value on the Likert scale. The corresponding factor loading in the index using factor analysis is relatively low, such that the question regarding the covering of the webcam has less weight than other questions (Appendix Table 3.4). Therefore, we conclude that our results from the additive index should be interpreted carefully, and a more precise index is needed to measure individual preferences for privacy on the internet. In Tables 3.3, 3.7 and 3.8 (Appendix), we still find partial support for H2 and H3. For these model specifications, stating the desire to lose weight increases the minimum $WTA_{details}$ roughly between ~€15.05 and ~€19.21 and decreasing the donated amount to the ILGA by one euro increases the minimum WTA_{DOT} between €0.32 and €0.39.

3.5 Conclusions

We conducted an experiment to investigate whether the privacy paradox is prevalent in an analogue classroom setting. We evaluate the WTA bids to disclose private information and DOT decisions using the BDM and relate the bids to privacy concerns, elicited via survey items. We find that most subjects are willing to forego considerable potential earnings in order to protect both private details and DOT decisions. Unlike previous studies (Benndorf and Normann, 2018; Plesch and Wolff, 2018), data disclosure in our experiment has no real life

equivalent. However, we think that the form of data disclosure we have chosen has resulted in participants having more similar expectations about how their data will be handled than if the data had been transferred to a physically more distant third party. In addition, the information disclosed in our study was only announced orally on one occasion and was not disclosed in written form as in Benndorf and Normann (2018) and Schudy and Utikal (2017). Therefore, the results of these studies and our results are only comparable to a limited extent. The prices for the willingness to reveal the personal details and the DOT decisions are significantly correlated with the stated privacy behavior when we use an additive index for privacy. However, this result is mainly driven by the individuals' stated decision to cover up the webcam of their laptop and should, therefore, not be overinterpreted. We find that delicate information, like stating the desire to lose weight and the DOT regarding the ILGA, significantly affects the willingness to reveal the information. A possible explanation for this behavior is that individuals try to avoid shameful exposure. Especially since our subjects were first semester students, they might have a particular interest in hiding possible embarrassing behavior in front of their new peers. However, we do not find conclusive and robust evidence for behavior which is not in line with the privacy paradox. Although we observe that individuals who state a higher preference for privacy in the questionnaire also demand a higher price for revealing their private information, this largely depends on the corresponding specification of our measure for privacy concerns. Future research would benefit from refined comprehensive indices measuring individual preferences for privacy on the internet.

Acknowledgements

We thank Matthias Greiff, the participants of the Clausthaler Ökonomisches Oberseminar (Clausthal University of Technology) and the participants of the Colloqium Recht & Ökonomie (University of Kassel).

3.6 Appendix

3.6.1 Experimental instructions

Welcome and thank you very much for participating!

Today you are taking part in an economic experiment. Please note that you can earn real money (more than 100 €) in this experiment and that the amount of money depends on your decisions in the experiment. At the end of the experiment three participants will be randomly selected, which are actually paid out. For the payment we can only consider questionnaires, which are completely filled out. Every participant of the experiment has the same chance to be randomly selected. Your personal decisions in the experiment have no influence on who is selected at the end. Those who are not selected will receive a small gift for the start of the semester as a thank-you. Details on the payment can be found on the yellow card with your identification number.

General rules of the experiment

We will now explain some general rules. Please follow these rules at all stages of the experiment. **From now on, please no longer communicate with the other participants.** This is to ensure that you make your decisions individually within the experiment and that your decisions cannot be observed by others. In addition, please always follow our instructions.

If you do not follow these rules, we will unfortunately have to exclude you from the experiment. In this case you lose the chance to be paid!

You may stop the experiment at any time and leave the auditorium. Please note, however, that we will not be able to pay you in this case. If you feel uncomfortable or already know that you will not be able to stay here for the next 30 minutes, please inform one of our team members.

Information on the experiment

This experiment is part of a study on the donation behavior of students. The whole experiment will require a maximum of 30 minutes of your time. During this time you will fill out various questionnaires. **It is essential that you enter your personal identification number, which you received from us at the beginning, at the top of each page.** Please note that there are no correct or incorrect answers to the questions we will ask you. Please always answer truthfully and according to your personal opinion. Please also note that the data collected in this experiment will be treated confidentially and will not be passed on to third parties.

You will never be deceived in the course of the experiment. All information provided by us is true.

If you have any questions, please raise your hand and wait quietly until somebody comes to you.

At the beginning we would like to ask you some questions:

1. **Some websites use special tools to identify returning users and to provide them with personalized information (such as advertisements). If a website offered you such a service, would you agree?**

In no case	Rather not	I don't know	Rather yes	In any case	I don't know / I don't want to answer
❒	❒	❒	❒	❒	❒

2. **How often do you use the private mode (incognito mode) of your web browser?**

Never	Rarely	Sometimes	Often	Allways	I don't know / I don't want to answer
❒	❒	❒	❒	❒	❒

3. **Do you use your real name on social networks (like Facebook)?**

Never	Rarely	Sometimes	Often	Allways	I don't know / I don't want to answer
❒	❒	❒	❒	❒	❒

4. **Do you allow apps on your smartphone, tablet or laptop to determine your location and record it if necessary?**

Never	Rarely	Sometimes	Often	Allways	I don't know / I don't want to answer
❒	❒	❒	❒	❒	❒

5. **Are you covering up your laptop's webcam?**

Never	Rarely	Sometimes	Often	Allways	I don't know / I don't want to answer
❒	❒	❒	❒	❒	❒

6. **How important is your privacy in general to you?**

Unimportant	Rather unimportant	Important	Very important	Extremely important	I don't know / I don't want to answer
❒	❒	❒	❒	❒	❒

7. **How important is the security of your personal data to you?**

Unimportant	Rather unimportant	Important	Very important	Extremely important	I don't know / I don't want to answer
❒	❒	❒	❒	❒	❒

Personal details

In this part of the experiment we ask you three questions. Simply choose the answer that best suits you.

1. Would you like to lose weight?
   ❒ yes ❒ no

2. Do you smoke?
   ❒ yes ❒ no

3. What was the average grade of your last school degree (e.g. Abitur)? _____

Donate-or-take-decisions

In this part of the experiment, you can make a decision about the distribution of an amount of money between yourself and one of three organizations. You start with 150 €. Every organization starts with 50 €. For each of the three organizations you can decide whether you donate between 0 € and 50 € to them or whether you take between 0 € and 50 € away from them.

The three organizations are the "KSV Hessen Kassel" (a local soccer club), the "ILGA" (International Lesbian, Gay, Bisexual, Trans and Intersex Association) and the "UNHCR" (UN Refugee Agency).

At the end of the experiment, three participants are randomly selected. For each of these participants we then randomly select an organization where the decisions of the particular participant will be implemented. Each organization is assigned to exactly one of the three randomly selected participants of the experiment.

Short summary:

Your endowment is 150 €.

The endowment of each organization is 50 €.

You can now make one of the following decisions for each organization:

> You donate an amount between 0 € and 50 € to the organization.
>
> **Or**
>
> You take an amount between 0 € and 50 € from the endowment of the organization. By this amount your payment increases, and the donation amount of the organization decreases accordingly.

Now please make a decision for each of the three organizations:

(Reminder: The decision will only be implemented if you and the organization affected are randomly selected.)

In case that the organization "KSV Hessen Kassel" (local soccer club) is randomly selected, I would like to ……

❒ take ❒ donate ❒ neither take nor donate money.

What amount would you like to donate or take (between 0 € and 50 €)?

_____ €

In case that the organization „ILGA“ (International Lesbian, Gay, Bisexual, Trans and Intersex Association) is randomly selected, I would like to ……

❒ take ❒ donate ❒ neither take nor donate money.

What amount would you like to donate or take (between 0 € and 50 €)?

_____ €

In case that the organization „UNHCR“ (UN Refugee Agency) is randomly selected, I would like to ……

❒ take ❒ donate ❒ neither take nor donate money.

What amount would you like to donate or take (between 0 € and 50 €)?

_____ €

After all participants have completed this part of the experiment, we collect all questionnaires. Afterwards the second part of the experiment begins. Please remain quietly seated at your place.

In this part of the experiment you can increase your payment amount. You can allow us to disclose information about you in return for payment, but you do not have to do so.

First of all, we will make you an offer in which, in the event of an implementation, your "Personal details" from the orange sheet (wish to lose weight, smoking, school grade) will be read out publicly and you will have to stand up during the reading so that all participants can see you.

The offer is structured as follows:

Below you can enter the price at which you are willing to give us the right to read out publicly your "Personal details" **if you are randomly selected and your decisions become relevant to your payment and that of the organization.**

Before the experiment, six prizes were written down by the experimenters and placed in separate envelopes. These prices are in the range of 0 € to 100 €. At the end of the experiment, two envelopes are randomly selected. If the price in the first envelope is lower than the price you specify below, your "personal details" **will not be publicly read** and you will **not receive any additional payment.**

If the price in the first envelope is higher than the price you specify, you will receive the **price from the envelope in addition to your existing payment. Your "personal details" will be read out publicly** and you will have to **stand up** during the reading so that all participants can see you.

Public reading of "personal details" (wish to lose weight, smoking, school grade)

Please tell us the price at which you are willing to give us the right to read out publicly your "Personal details" while you are standing.

(In no case we will pay you more than 100 € for this right. If you enter more than 100 €, your "Personal details" will not be read out publicly under any circumstances.)

Minimum price at which my "Personal details" can be read out publicly while I stand up for this:

_____ €

In addition, we make you an offer in which, in case of an implementation, your "Donate-or-take-decisions" from the orange sheet are read out publicly and you have to stand up during the reading so that all participants can see you.

The offer is structured as follows:

Below you can enter the price at which you are willing to give us the right to read out publicly your "Donate-or-take-decisions" **if you are randomly selected and your decisions become relevant to your payment and that of the organization.**

Before the experiment, six prizes were written down by the experimenters and placed in separate envelopes. These prices are in the range of 0 € to 100 €. At the end of the experiment, two envelopes are randomly selected. If the price in the first envelope is lower than the price you specify below, your "Donate-or-take-decisions" **will not be publicly read** and you will **not receive any additional payment.**

If the price in the first envelope is higher than the price you specify, you will receive the **price from the envelope in addition to your existing payment. Your "Donate-or-take-decisions" will be read out publicly** and you will have to **stand up** during the reading so that all participants can see you.

Public reading of „Donate-or-take-decisions"

Please tell us the price at which you are willing to give us the right to read out publicly your "Donate-or-take-decisions" while you are standing.

(In no case we will pay you more than 100 € for this right. If you enter more than 100 €, your "Donate-or-take-decisions" will not be read out publicly under any circumstances.)

Minimum price at which my "Donate-or-take-decisions" can be read out publicly while I stand up for this:

_____ €

Please go to the next page for some final questions.

We would be grateful if you could (voluntarily) answer the following question: What were the reasons for your decisions on the two previous pages?

At this point you can give us general feedback on the experiment:

Demographic data

Your gender: ❒ male ❒ female ❒ other / no speficiation

Your age: _____

Thank you very much for your participation!

3.6.2 Tables

	(1) $WTA_{details}$	(2) $WTA_{details}$	(3) $WTA_{details}$	(4) WTA_{DOT}	(5) WTA_{DOT}	(6) WTA_{DOT}
privacy	12.38**	11.48**	11.76**	12.15*	12.39*	10.19
	(5.80)	(5.74)	(5.89)	(6.78)	(6.69)	(6.82)
weight		15.05*	17.83**			
		(8.41)	(8.74)			
smoke		7.98	7.28			
		(10.28)	(10.37)			
grade		−2.81	−2.27			
		(7.13)	(7.19)			
KSV					0.13	0.14
					(0.14)	(0.14)
ILGA					−0.34**	−0.32*
					(0.17)	(0.18)
UNHCR					0.06	0.07
					(0.20)	(0.20)
age			−0.57			3.55*
			(1.77)			(2.11)
female			−4.45			7.81
			(8.97)			(10.26)
session	10.99	12.44	11.36	19.91**	17.20*	17.93*
	(8.41)	(8.73)	(9.43)	(9.72)	(9.75)	(10.52)
Constant	10.08	11.32	22.83	8.26	13.68	−57.99
	(16.95)	(26.36)	(44.55)	(19.81)	(20.17)	(46.31)
sigma	38.95***	38.06***	38.09***	44.66***	43.64***	43.27***
	(3.28)	(3.20)	(3.25)	(3.92)	(3.83)	(3.85)
Observations	94	94	92	94	94	92
pseudo R^2	0.01	0.01	0.01	0.01	0.02	0.02

Table 3.3: Tobit regressions of WTA on self-stated privacy behavior (additive a1-a5) and information at stake.

Notes. Robust standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. Out of 99 individuals, we lose five in the first four model specifications since some participants chose "I don't know/don't want to answer" for some of the privacy questions. Additionally, we lose two more observations in models (3) and (6) because two participants did not report their age.

Variable	Factor	Uniqueness
a1	0.49	0.76
a2	0.26	0.93
a3	0.34	0.88
a4	0.53	0.72
a5	0.12	0.98
a6	0.67	0.55
a7	0.58	0.66
eigenvalue	1.51	
alpha	0.54	

Table 3.4: Factor loadings using all privacy questions.

Variable	Factor	Uniqueness
a1	0.49	0.76
a4	0.51	0.74
a6	0.65	0.58
a7	0.58	0.66
eigenvalue	1.25	
alpha	0.65	

Table 3.5: Factor loadings using factor with highest alpha.

Indices	Factor Analysis (a1–a7)	Factor Analysis (a1, a3, a6, a7)	Additive (a1–a7)	Additive (a1–a5)
factor analysis (a1-a7)	1.00			
factor analysis (a1, a3, a6, a7)	0.98***	1.00		
additive (a1–a7)	0.89***	0.81***	1.00	
additive (a1–a5)	0.73***	0.62***	0.95***	1.00

Table 3.6: Correlation between different privacy concern specifications.

Notes. Robust standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.

	(1) $WTA_{details}$	(2) $WTA_{details}$	(3) $WTA_{details}$	(4) WTA_{DOT}	(5) WTA_{DOT}	(6) WTA_{DOT}
privacy	0.96	4.70	5.02	6.51	8.79	1.92
	(23.03)	(22.27)	(22.56)	(24.97)	(24.96)	(24.68)
weight		16.94*	19.21**			
		(8.71)	(9.30)			
smoke		7.74	7.29			
		(10.66)	(10.84)			
grade		−4.44	−3.80			
		(7.04)	(7.11)			
KSV					0.11	0.11
					(0.14)	(0.14)
ILGA					−0.35**	−0.32*
					(0.17)	(0.18)
UNHCR					0.09	0.09
					(0.23)	(0.23)
age			−0.33			3.85*
			(1.74)			(1.96)
female			−1.65			10.69
			(8.96)			(9.61)
session	12.79	13.51	13.55	21.42**	18.38*	19.98*
	(8.47)	(8.64)	(9.16)	(9.48)	(9.62)	(10.39)
Constant	43.01***	43.61*	48.75	37.81**	42.13***	−40.17
	(13.20)	(24.00)	(40.67)	(14.83)	(15.93)	(40.34)
sigma	40.17***	39.09***	39.14***	45.48***	44.47***	43.83***
	(2.98)	(3.04)	(3.14)	(3.25)	(3.28)	(3.30)
Observations	94	94	92	94	94	92
pseudo R^2	0.00	0.01	0.01	0.01	0.01	0.02

Table 3.7: Tobit regressions of WTA on self-stated privacy behavior (factor analysis a1-a7) and information at stake.

Notes. Robust standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. Out of 99 individuals, we lose five in the first four model specifications since some participants chose "I don't know/don't want to answer" for some of the privacy questions. Additionally, we lose two more observations in models (3) and (6) because two participants did not report their age.

	(1) $WTA_{details}$	(2) $WTA_{details}$	(3) $WTA_{details}$	(4) WTA_{DOT}	(5) WTA_{DOT}	(6) WTA_{DOT}
privacy	−5.82	−1.52	−1.34	5.12	5.65	−0.90
	(22.91)	(22.35)	(22.67)	(24.18)	(24.03)	(23.77)
weight		17.03**	19.10**			
		(8.37)	(8.97)			
smoke		5.40	5.04			
		(10.62)	(10.80)			
grade		−4.16	−3.58			
		(6.71)	(6.86)			
KSV					0.14	0.14
					(0.14)	(0.14)
ILGA					−0.39**	−0.36**
					(0.17)	(0.18)
UNHCR					0.09	0.10
					(0.22)	(0.22)
age			−0.41			3.61*
			(1.71)			(1.90)
female			−1.13			12.42
			(8.82)			(9.23)
session	17.04**	17.65**	17.99**	23.35**	20.22**	22.45**
	(8.19)	(8.45)	(9.02)	(9.11)	(9.23)	(9.98)
Constant	44.13***	43.65*	50.18	37.92**	43.43***	−35.20
	(14.17)	(23.75)	(40.27)	(15.46)	(16.18)	(39.76)
sigma	40.22***	39.17***	39.22***	45.04***	43.82***	43.10***
	(2.88)	(2.93)	(3.02)	(3.11)	(3.13)	(3.16)
Observations	99	99	97	99	99	97
pseudo R^2	0.00	0.01	0.01	0.01	0.01	0.02

Table 3.8: Tobit regressions of WTA on self-stated privacy behavior (factor analysis a1, a3, a6, a7) and information at stake.

Notes. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. We lose two observations in models (3) and (6) because two participants did not report their age.

	(1) $WTA_{details}$	(2) WTA_{DOT}
a1	−0.91	−0.70
	(3.58)	(3.54)
a2	0.22	−0.47
	(3.22)	(3.39)
a3	2.71	0.03
	(2.93)	(3.02)
a4	−1.39	1.32
	(3.70)	(3.69)
a5	8.38***	7.18***
	(1.94)	(2.13)
Constant	25.95*	29.28*
	(15.47)	(17.28)
Observations	94	94
F	4.28	2.35
R^2	0.19	0.12

Table 3.9: OLS regression of WTA on privacy behavior.

Notes. Robust standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.

4. Employer review platforms – do the rating environment and platform design affect the informativeness of reviews? Theory, evidence, and suggestions

This chapter has been published as:

Cloos, J. (2021). Employer Review Platforms–Do the Rating Environment and Platform Design affect the Informativeness of Reviews? Theory, Evidence, and Suggestions. *mrev management revue*, *32*(3), 152-181.

This paper can be downloaded via

doi.org/10.5771/0935-9915-2021-3-152

Abstract: Online employer review platforms (ERPs) enable employees to evaluate their current and former companies anonymously online. Job-seekers can use the aggregated reviews to obtain information about potentially attractive companies and thus limit the number of suitable companies. However, the matching process between job-seekers and companies can only be effective if the information provided on ERPs is representative and can be trusted. This paper investigates specific characteristics of ERPs using the two large ERPs *Kununu* and *Glassdoor* as examples. It is argued that the ERP environment is very different from the well-known and -studied reputation system environment of online marketplaces, and that specific factors can potentially bias reviews on ERPs. Based on a new data set containing the *Kununu* and *Glassdoor* reviews of 114 major German employers, it is analyzed if and how design aspects of ERPs and other specific factors affect reviews. Results show that overall (and industry specific), average review scores on *Kununu* and *Glassdoor* differ significantly from each other. Further results indicate that factors such as employees' awareness of their impact on a company's reputation also affect reviews. Suggestions are made on how ERPs could reduce the influence of these factors in order to present the aggregated information more effectively. (JEL: C81, M50, M51, M54)

Keywords: employer reviews, reputation, work standards, rating systems, online marketplaces

4.1 Introduction

The success story of e-commerce is closely linked to the successful establishment of various online rating (or reputation) systems. These rating systems enable users of online marketplaces such as *eBay*, *Amazon*, or *Airbnb* to assess the trustworthiness of other users and the quality of the products and services offered. In this way, online marketplaces can carry out transactions that would not have been possible without the existence of well-functioning rating systems that minimize transaction costs (Luca, 2017; Tadelis, 2016).

In the course of the steady growth and continuous improvement of online marketplaces, the benefits of online-based rating systems for human resource management (HRM) were also recognized. Accordingly, in the mid-2000s, the first online employer review platforms (ERPs) were launched. In the years before, several platforms such as *Monster*, *JobScout24*, or *StepStone* had already been established for job placement and to maintain business contacts via the internet, enabling employees and companies to get in touch with each other (Grund, 2006).

ERPs like the Austrian company *Kununu* and the US-company *Glassdoor* have greatly expanded their range of services since the founding years and are recording continuous growth. The information obtained via ERPs differs from information provided by the companies themselves, e.g., at job fairs or official websites, and reflects a broader spectrum of individual opinions. Since the information voluntarily provided on ERPs can reduce information asymmetries between employees and companies during the process of finding a job, the quality of employee-job matches can be improved.[21] ERPs also have the potential to enhance the relationship between a company and its current employees. As a result of the digital transformation, employees in many companies are faced with changes in their work-life setup (Schwarzmüller et al., 2018). Companies can get important feedback on how to successfully manage these changes through the information provided on ERPs.

Since each individual review only reflects the subjective judgement of a single employee and thus has limited information content, a higher number of reviews allows to draw better conclusions regarding the actual quality of a company. By aggregating as many individual reviews as possible, ERPs promise greater transparency on the labor market. It seems likely that after a successful job search, employees will recommend ERPs and use them repeatedly, especially the more accurately the company information provided on ERPs corresponds to the

[21] Grund (2006) discusses the implications of matching theory for the employee recruitment process via the internet in detail.

actual conditions (experienced) in a company. However, it remains an open question whether and how the concrete design of an ERP influences the reviews and how representative the reviews really are for a company's workforce as a whole.

When evaluating their current or former company on an ERP, employees often provide more sensitive information compared to rating completed transactions or purchased products on *eBay* or *Amazon*. For online marketplaces, several studies identify review influencing factors such as the reciprocity between the buyer and the seller (Ye et al., 2014; Bolton et al., 2013), the gender of the reviewer (Craciun and Moore, 2019) or emotional expressions in reviews (Kim and Gupta, 2012). A further but yet uninvestigated question is therefore whether specific factors on ERPs may potentially bias employees' aggregate reviews. The aim of this article is to fill this research gap by analyzing and comparing the average review scores on *Kununu* and *Glassdoor* systematically for different industries and, in the case of *Kununu*, for different subgroups of employees and reviews. The results show that the average review scores on *Glassdoor* are significantly better than those on *Kununu*. It is argued that these discrepancies result from differences in ERP designs which lead to different perceived levels of anonymity and differences in self-selection. It is further theoretically explained and, wherever possible, empirically demonstrated that the time of posting a review, socially influenced preferences, and employees' awareness of their impact on a company's reputation can affect the reviews of different subgroups.

This paper has the following contributions: First, ERP operators gain insights that can help to improve the existing ERP designs and thereby attract further customers in form of companies and employees. Second, employees, job-seekers, and other stakeholders can gain information on aspects which should be considered when interpreting the contents of ERPs. And third, as this is one of the first papers that examines the ERP environment in detail, I hope that this paper provides a stepping stone for future research on particular characteristics of ERPs.

The remainder of the paper is organized as follows. Section 4.2 gives a brief summary of related literature on ERPs. Section 4.3 addresses the main differences between the rating systems of online marketplaces and ERPs. Section 4.4 describes the review systems of *Kununu* and *Glassdoor*, the data set, and addresses the ERP range and usage behavior of employees and companies. In section 4.5, specific factors that can influence ERP reviews are examined both theoretically and empirically. Suggestions on how to minimize the impact of these factors are presented in section 4.6. The last section concludes by pointing out a number of limitations and outlining a research agenda.

4.2 Literature on employer review platforms

In recent years, ERPs have been increasingly used not only by employees and companies, but also for research purposes. This section aims to provide a brief overview of these studies and some exemplary findings, without any claim to completeness. The first subsection presents studies that use information provided on ERPs for a variety of research objectives. Thereafter, studies examining explicit design features of ERPs or addressing questions regarding the reliability of ERP reviews are presented.

4.2.1 Using ERP data for research

So far, there is a small but growing number of studies that use the information available on ERPs as a data source for different research questions. The majority of these studies (Marinescu et al., 2021; Huang et al., 2020; Dabirian et al., 2016; Luo et al., 2016; Moniz, 2015; Moniz and de Jong, 2014) rely on data from *Glassdoor* but other recently published studies (Hoon et al., 2019; Kollitz et al., 2019; Könsgen et al., 2018; Abel et al., 2017) also use *Kununu* as their data source. A large number of studies (Könsgen et al., 2018; Abel et al., 2017; Luo et al., 2016; Dabirian et al., 2016; Moniz, 2015; Moniz and de Jong, 2014) also use text mining tools in order to categorize the reviews in regard to their linguistic content.

The results of Huang et al.'s (2020) study show that employees' business outlooks collected from *Glassdoor* are well suited to predict the future operating performance of companies. The studies of Luo et al. (2016), Moniz (2015) and Huang et al. (2015) examine the relation between ERP contents and the financial performance of companies. The results of all three studies indicate that there is a positive correlation between a company's review score on ERPs and Tobin's q. Following the "Dieselgate" scandal, Hoon et al. (2019) examine more than 1,000 *Kununu* reviews of *Volkswagen* employees and find that they showed no increasingly destructive voice behavior towards their company after the scandal, but that the amount of constructive voice behavior decreased. Kollitz et al. (2019) use company review scores from *Kununu* as an external measure of employer reputation in a study on the recruitment strategies of family businesses. The authors find that below-average recruitment practices predict poor employee ratings on *Kununu*.

4.2.2 Studies on the validity of reviews and specific ERP designs

The scientific literature dealing with the specific characteristics of ERPs is limited. Marinescu et al. (2021) refer to the fact that the existing literature has not yet investigated to what extent the online review behavior of employees differs from the relatively well studied

online review behavior of consumers (e.g., Dorner et al., 2020; Filippas et al., 2018; Bolton et al., 2013).

By comparing the *Glassdoor* review scores of US federal agencies' employees, Landers et al. (2019) examine the construct validity of the reviews on *Glassdoor*. The authors find that the general job satisfaction information provided on *Glassdoor* can be considered valid as the values from *Glassdoor* correlate moderately with the values from the official survey.

The only study that examines a concrete design feature of an ERP is Marinescu et al. (2021). *Glassdoor* requires its users to provide work-related information in return for unrestricted access to the available information. According to *Glassdoor* this "Give to get policy" is intended to ensure that the written reviews reflect a wide range of opinions. The results of Marinescu et al. (2021) show that the "Give to get policy" caused a slight but significant increase of 2.6 (2.9%) in the proportion of rather moderate 3 (4) star reviews compared to voluntary reviews. The share of the worst (best) 1 (5) star ratings decreased significantly by 3.6 (2.1%). Following the interpretation of Marinescu et al. (2021), the non-monetary incentive to provide a review results in reviews reflecting a more representative picture of employee opinions in the aggregate.

4.3 Differences between the rating systems of ERPs and online marketplaces

As previously mentioned, the scientific literature on rating systems of online marketplaces identifies various factors that can influence the contents of reviews. Addressing existing differences between the rating systems of online marketplaces and the rating systems of ERPs is helpful in order to determine whether and which of these factors are also relevant for the reviews on ERPs. Following the discussion in this section, section 4.5 turns to examining factors that can specifically affect reviews on ERPs in more detail.

4.3.1 What is the subject of the review?

On internet platforms like *eBay*, *Amazon*, or *Yelp*, users evaluate products purchased or services. In general, one-time transactions are rated. Users of these platforms therefore have little reason to assume that they will suffer negative consequences as a result of their review. In contrast, on *Kununu* and *Glassdoor*, employees evaluate an organization to which they actually belong or have belonged to in the past. The submission of a review on an ERP obviously has a stronger potential to cause negative consequences that affect the reviewing person afterwards. In the case of current employees, the organization in question has a direct influence on the

economic situation of the respective employee. In this relatively sensitive review environment, other factors can have an impact on the provided contents of reviews compared to online marketplaces.

Users of online marketplaces in general rate a relatively homogeneous product or service. Although it is possible that different attributes may be taken into account when posting a review, in most cases, the number of the attributes is relatively small. In contrast, the users of ERPs evaluate a complex company. In a company, employees work in different business units, have different colleagues and supervisors, and deal with different tasks. The subject area evaluated on ERPs is therefore much more heterogeneous than on online marketplaces.

4.3.2 What are the motives for providing a review?

A greater similarity between the rating systems of ERPs and online marketplaces can be found in terms of the potential motives that encourage users to submit a review. First of all, it is plausible that people who provide a review are guided by the motive to contribute to a public good in both review environments. The users of online marketplaces and ERPs are therefore aware that they benefit from a large pool of reviews when making their own decisions and want to contribute to this pool in order to increase the amount of information available.

However, on online marketplaces such as *eBay* (buyers and sellers) or *Airbnb* (hosts and guests), reviews are provided by both market sides. Every user who provides a review has an own account with a nickname and can also receive reviews from transaction partners. Due to these accounts, the users of online marketplaces are often identifiable at least in terms of their gender or skin color. Several studies (Cui et al., 2020; Edelman et al., 2017; Ayres et al., 2015; Doleac and Stein, 2013; Nunley et al., 2011) show that this partial identifiability allows for discrimination. On *Amazon,* users have accounts as well which can incentivize them to provide reviews that are as informative as possible. Every user can mark reviews from other users that she regards as helpful. Users whose reviews have been particularly often marked as helpful can receive special benefits from *Amazon* by being included in an exclusive club of product testers (Dorner et al., 2020). Since employees provide reviews on ERPs on a completely anonymous basis (accounts are invisible for other users), problems such as discrimination do not play a role. Nevertheless, the missing possibility to mark existing reviews as helpful also reduces the incentive to provide highly informative reviews.

Although ERPs pursue economic interests, they can, in the broader sense be assigned to a commons-based peer production environment (Benkler, 2006). Algan et al. (2013) use the

example of the online encyclopedia *Wikipedia* to investigate motives that tempt individuals to engage in such commons-based peer production environments. Using an online experiment in which *Wikipedia* authors acted as test participants, the authors show that the number of contents that the test participants had contributed to *Wikipedia* was strongly related to their preference for reciprocal exchange, their social image interest, and their altruistic preferences. Therefore, an individual's motivation to submit a review on an ERP can also be explained by intrinsic motivation (see e.g., Poch and Martin, 2015; Bitzer et al., 2007; Tedjamulia et al.; 2005; Kreps, 1997) and reciprocity (see e.g., Jochims, 2016; Fehr and Gächter, 1998).

4.4 ERP characteristics, data set, and descriptive statistics

4.4.1 Kununus and Glassdoors rating systems

On *Kununu* and *Glassdoor*, employees can rate companies in different categories by using a five star scale. On *Kununu*, stars can be awarded in 13 categories such as working atmosphere, supervisor behavior and working conditions. On *Glassdoor*, in addition to the overall rating, stars can be awarded in five other categories such as work/life balance and career opportunities. On both ERPs, employees can describe individual experiences in text comments and indicate whether they would recommend their current or former company. Reviews can be submitted by active and former employees as well as trainees, interns and applicants.

The submission of reviews takes place anonymously. Employees who wish to leave a review must, however, register with a valid E-mail address on the respective ERP. *Kununu* and *Glassdoor* have established various technical and manual testing procedures to ensure that the reviews are written by actual employees and meet the codes of conduct. Furthermore, both ERPs promise that they never delete or change the contents of reviews as long as these reviews met the codes of conduct.[22] When writing a review on *Kununu* and *Glassdoor*, various information may be deliberately omitted, especially if a piece of information allows identification. A difference between the two ERPs is, that the location of the company to be reviewed must always be indicated on *Kununu*, while on *Glassdoor* this information can be deliberately omitted. On both ERPs, reviews can be filtered by different employee subgroups,

22 See videos (in German) on *Kununus'* review control system (https://kununugmbh.zendesk.com/hc/de/articles/115004243929-Wie-sorgt-kununu-f%C3%BCr-echte-Bewertungen-wie-funktioniert-die-Bewertungskontrolle-) and codes of conduct (https://kununugmbh.zendesk.com/hc/de/articles/115004235245-Warum-gibt-es-auf-kununu-Regeln-wie-lauten-diese-) and *Glassdoors'* community guidelines (https://help.glassdoor.com/article/Community-Guidelines/en_US) (all three sources accessed June 06, 2020).

review score, and on *Kununu*, by the time period (i.e. reviews written in the last month, the last 6 or 12 months).

On both ERPs, companies can choose between free and paid company profiles, whereby the paid variants include far more options. Companies being active on the respective ERP are marked accordingly.

Company-related information is freely accessible to every user on *Kununu*. In addition to all reviews, employer responses to reviews and company profiles (if existent) can be viewed for free. *Glassdoor* users are initially presented with only a limited amount of user generated content for each company. While users can, for example, see the average review score, only a limited number of reviews can be viewed in detail.

Kununu's unique characteristic is the connection to the career-oriented social networking service *XING*, established in January 2013. The connection to *XING* appears quite advantageous for *Kununu*. Employees who have a *XING* profile but have not yet been active on an ERP will be approached by *XING* to use *Kununu*.

4.4.2 Data set

This section explains the data set, which is used in the subsequent section to present descriptive statistics on the ERP range and the ERP usage behavior of companies and employees in Germany. In section 4.5, the data set is used to perform more detailed statistical tests and a regression analysis that investigates potential bias factors that may affect the reviews on ERPs. The aim of these investigations is to gain deeper insights into the specific characteristics of ERPs, which are not covered by the literature presented in section 4.2.

The dataset contains several indicators from *Glassdoor* and *Kununu* and key figures for 114 companies. The companies were selected on the basis of the biennial report of the German Monopolies Commission (Monopolkommission, 2018). In terms of domestic net product, the report contains the largest companies in Germany for the reporting years 2016 and 2014. The partial geographical limitation to the largest European economy, Germany, allows a detailed analysis of the two ERPs and at the same time ensures a sufficiently large data basis.

The data set contains the number of employees in Germany for the year 2016 (2014) for 108 (97) companies. In 2016, these companies thus employed at least 3.636.987 people, which

corresponds to more than 8.28% of the employed German residents during that time period.[23] In order to enable a uniform comparison of the ERPs and companies, all data from *Kununu* and *Glassdoor* were collected within one week (April 6-9, 2020).

For *Kununu*, the total number of reviews, the number of reviews for the last 12 months, the average review score, the recommendation rate, the number of employer responses, and the number of active months on *Kununu* were collected for each company. In order to enable a detailed analysis, data on the number of reviews, the average review score, and the recommendation rate were also collected for the following subsets: current and former employees, executives and non-executives, and reviews with and without employer responses. Due to limited filter functions for *Glassdoor*, only the total number of reviews, the average review score, the recommendation rate and the number of active months on *Glassdoor* were collected for each company. Since the recommendation rates correlate strongly with the average review scores (*Kununu*: $r = 0.92$, p–value < 0.001; *Glassdoor*: $r = 0.83$, p-value < 0.001), only the average review scores are used in the further analyses.

For both ERPs, it was also recorded whether the companies were marked as active employers on the respective ERP. The companies were assigned to one of eleven (clustered) industries based on their indicated economic sector (see Appendix section 4.8).[24] For all figures, tables, regressions, and significance tests in this paper, only those companies were considered for which at least 10 reviews had been submitted on the respective ERP.

4.4.3 ERP range and usage of companies and employees in Germany

With approximately 60 million reviews, salary reports, and insights on more than one million companies and 50 million different visitors per month, *Glassdoor* is currently (May, 2020) one of the world's largest ERPs.[25] More than 7,000 companies are customers of *Glassdoor* and make

[23] According to the Federal Statistical Office of Germany (2019), there were 43,900,000 employed German residents in December 2016.

[24] Throughout this paper, due to lack of space, only the first-mentioned industries of the clustered industries are mentioned in figures, tables, and the text.

[25] It should be noted that the job search engine *Indeed* has the largest number of company reviews for worldwide locations. However, since *Indeed's* number of reviews for companies in Germany is much lower and the filter functions are less advanced than on *Kununu* and *Glassdoor*, this paper refrains from taking a closer look at *Indeed*.

use of the available recruitment and advertising opportunities.[26] In 2015, the platform was already actively used by 433 (87%) of the Fortune 500 companies (Barnes et al., 2015).

In German-speaking countries, *Kununu* is the largest ERP with more than 4.1 million reviews on more than 946,000 companies.[27] As of May 2020, *Kununu* has been active in German-speaking countries for more than 12 years and has therefore been present in this region much longer compared to *Glassdoor* (since January 2015). In addition to *Kununu* and *Glassdoor*, there are further ERPs in German-speaking countries which are not taken into account in this paper due to the much lower user numbers (see Reuter and Junge, 2017).

In the following, several indicators collected directly from both ERPs for selected companies and their employees are presented by using the data set. In this way, a more detailed picture of the current ERP usage behavior of employees and companies in Germany is provided. As mentioned above, for reviews on *Kununu* there is a need to indicate a company's location. As this is not the case for *Glassdoor*, for the companies included in the data set only a comparatively small proportion of the total reviews can be clearly attributed to company locations in Germany.

Figure 4.1 shows whether the companies in the dataset are marked as active employers on *Kununu* and/or *Glassdoor*. Only 9 out of 114 companies (7.89%) are not active on either of the two ERPs. Of the 114 companies considered, 88 (77.19%) are active on *Kununu* and 79 (69.30%) on *Glassdoor*.

[26] The information was extracted from the 'about us' section on the *Glassdoor* website (https://www.glassdoor.com/about-us/, https://www.glassdoor.com/about-us/recruit-holdings-announces-completion-of-glassdoor-acquisition/, both accessed June 06, 2020).

[27] The information was extracted from the main section on the *Kununu* website (https://www.kununu.com/, accessed June 06, 2020).

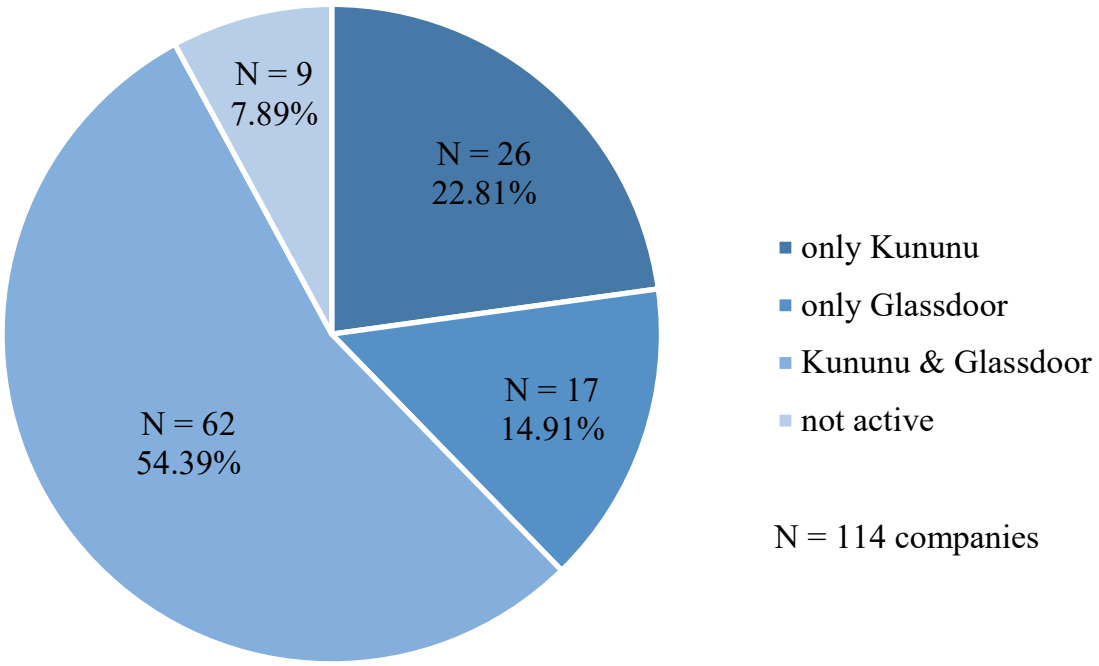


Figure 4.1: ERP activities of the companies.

Table 4.1 presents values of several indicators on *Kununu* and *Glassdoor* ordered by company sizes. The number of reviews for German locations is considerably higher for *Kununu*. As it is not mandatory to indicate the location of the company when submitting a review on *Glassdoor*, Table 4.1 also presents the numbers of reviews relating to worldwide locations which are, on average, much higher. For the companies considered, less than 5% of the reviews on *Glassdoor* can be clearly identified with locations in Germany. Even when considering all reviews in German (language), the number of these reviews accounts, on average, for less than 10% of the worldwide reviews for *Glassdoor* (location indicated and not indicated).

On average, less than 10% of reviews on *Kununu* received responses from the respective companies. Moreover, more than 40% of the companies use the response function only to a very limited extent. A closer comparison of the average review scores of reviews with and without employer responses shows that the companies use the response function primarily to respond to reviews with a below average star rating. Considering companies that have responded on at least 10 reviews, the mean review score is significantly worse (unpaired t-test (unequal variances), p-value < 0.01, N = 65 companies) for reviews with responses (mean = 3.33, sd = 0.54) than for reviews without responses (mean = 3.64, sd = 0.26).

Company Size (in 2016)		< 10.000	10.000-19.999	20.000 – 49.999	≥ 50.000	All Sizes
Companies	N	35	39	20	20	114
	%	30.70	34.21	17.54	17.54	100.00
			Kununu			
Reviews (full period)	Mean	453.80	547.21	861.90	2123.75	850.32
	Median	374.00	463.00	880.50	1720.00	534.00
	(sd)	(331.32)	(430.75)	(535.29)	(1567.73)	(961.41)
	N	35	39	20	20	114
Reviews (last 12 months)	Mean	96.97	116.38	172.35	525.90	192.08
	Median	56.00	108.00	147.50	386.50	116.50
	(sd)	(75.90)	(86.08)	(250.95)	(432.43)	(251.92)
	N	35	39	20	20	114
Employer Responses	Mean	50.94	48.54	75.50	195.80	79.84
	Median	8	19	26.50	66.50	23.50
	(sd)	(94.95)	(93.85)	(96.85)	(237.35)	(140.29)
	N	35	39	20	20	114
Companies with <10 Employer Responses	N	19	16	8	6	49
	% per Group	54.29	41.03	40.00	30.00	42.98
Companies marked as Active Employers	N	25	32	14	17	88
	% per Group	71.43	82.05	70.00	85.00	77.19
			Glassdoor			
Reviews (Germany)	Mean	62.80	83.24	95.60	270.10	112.43
	Median	30.00	37.00	56.00	224.00	46.50
	(sd)	(101.82)	(228.79)	(95.14)	(190.40)	(183.03)
	N	35	37	20	20	112
Reviews (language: German)	Mean	142.66	149.24	162.70	488.15	210.11
	Median	68.00	77.00	98.00	347.00	85.00
	(sd)	(272.28)	(256.99)	(162.64)	(381.97)	(301.05)
	N	35	37	20	20	112
Reviews (Worldwide)	Mean	3223.14	2918.61	1154.80	2023.95	2542.41
	Median	216.00	287.50	240.00	1230.00	417.00
	(sd)	(7407.77)	(8963.95)	(1762.12)	(2334.07)	(6726.01)
	N	35	38	20	20	113
Companies marked as Active Employers	N	22	25	14	18	79
	% per Group	62.86	64.10	70.00	90.00	69.30

Table 4.1: ERP usage characteristics for Kununu and Glassdoor ordered by company size (employees in Germany in 2016).

4.5 Possible bias factors affecting reviews on ERPs

Section 4.3 has emphasized that the review environment of ERPs differs considerably from the review environment of online marketplaces, especially with regard to the subject of the review. Below, four factors that can bias the (aggregate) reviews on ERPs are identified and examined. The sequence of these factors is chosen with regard to their presumed impact on average reviews where those factors considered to have the biggest influence are examined first.

4.5.1 Perceived level of anonymity

ERPs state that they are concerned with ensuring the highest possible degree of anonymity for their users. These efforts are quite understandable from a scientific point of view. For example,

Brutus and Derayeh (2002) and Antonioni (1994) show that a high degree of anonymity encourages honest employer evaluations.

Even if ERPs try to guarantee a high degree of anonymity, it seems questionable whether this is perceived by ERP users accordingly. In particular, employees who work for relatively small companies may fear that they could be identified through the submission of a review.[28]

If employees have the ambition to write a review that is as informative as possible, they often not only award stars in different categories, but also write text commentaries in which they describe their individual job-related experiences. The submitted experiences could potentially harm the rated company.[29] If an employee is concerned that the publication of her work-related experience will enable her identification and could cause negative consequences for the company, she could deliberately limit her review to more positive experiences. As Cloos et al. (2019) experimentally show, the willingness to accept disclosure of information to other people depends strongly on the concrete content of this information. Further, the results of several experimental studies (Cloos et al., 2019; Benndorf and Normann, 2018; Schudy and Utikal, 2017) suggest that between 10 and 20% of participants generally refuse to share their private information with others. Transferred to ERPs, this implies that employees with concerns about the consequences of their shared work-related experiences could deliberately refrain from writing text commentaries and/or specifying the location of their employer. In the case of too strong concerns, they could also decide completely against the submission of a review.

As mentioned earlier, employees who submit a review on *Glassdoor* can choose not to disclose their company's location and obviously often decide against a disclosure. As shown in Table 4.1, the number of reviews on *Glassdoor* written in German is almost twice as high as the number of reviews written in German and additionally indicating a company location.[30]

[28] Media reports in which German employers could clearly assign negative reviews to individual employees can be found online. In one case, this even resulted in dismissal (see (in German): https://www.waz.de/wirtschaft/noten-fuer-den-chef-kann-man-job-bewertungsportalen-trauen-id214021731.html, https://www.handelsblatt.com/unternehmen/beruf-und-buero/the_shift/arbeitgeber-bewertung-im-netz-kantinenessen-lecker-kollegen-nett-chef-bloed/21086920.html, both accessed June 06, 2020).

[29] See Pfeffer et al. (2014) on negative word-of-mouth dynamics in social media networks.

[30] It is of course possible that some of the reviews in German refer to company locations in Austria, Switzerland, and other countries. However, it is assumed here that these reviews represent only a negligible percentage share, since most of the locations of the companies in the data set are in Germany, and Germany also has considerably more employed inhabitants than Austria and Switzerland.

Unlike on *Glassdoor*, employees on *Kununu* are required to indicate the location of their company when providing a review. For reviews with an indicated company location it therefore seems likely that these reviews reflect a wider range of opinions when provided on *Kununu* then when provided on *Glassdoor*. In order to test whether the perceived level of anonymity can have an influence, *Kununu's* and *Glassdoor's* average review scores for reviews with an indicated company location (in Germany) are compared below.

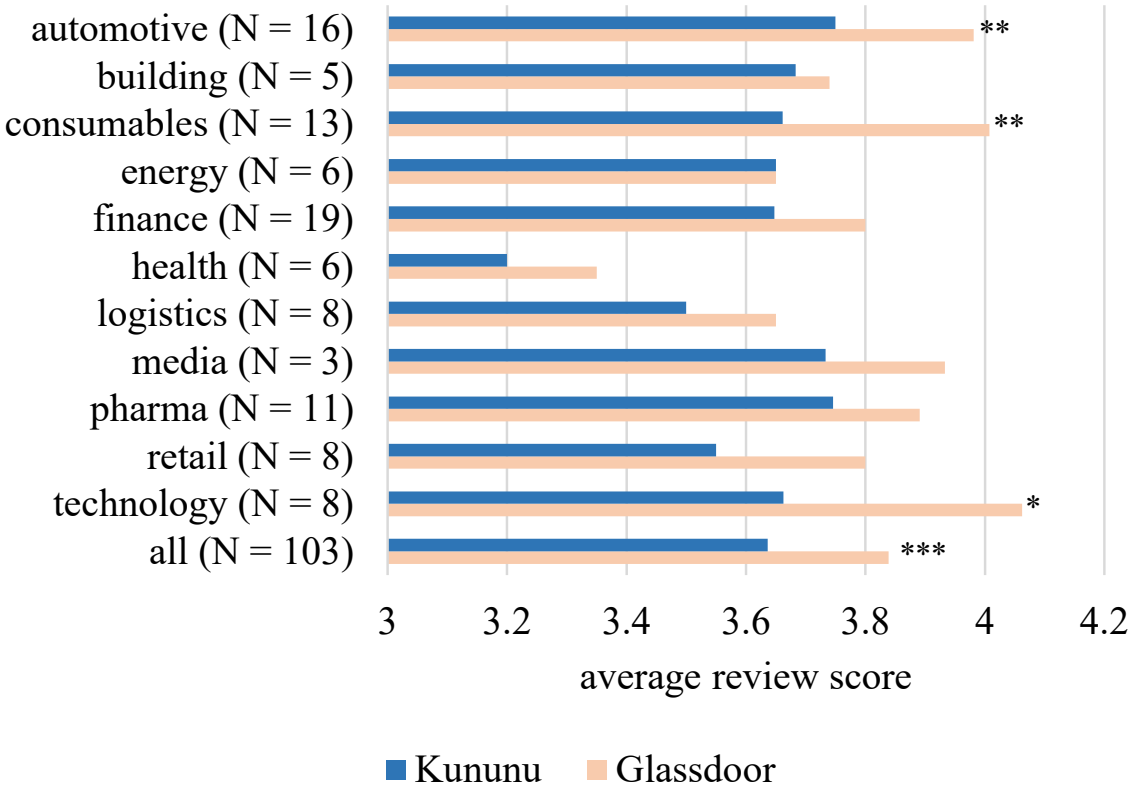


Figure 4.2: Mean review scores for Kununu and Glassdoor ordered by industry.

Notes. Unpaired t-tests (unequal variances) for the following industries: automotive, building, finance, health, logistics, pharma, all. Unpaired t-tests (equal variances) for the following industries: consumables, energy, retail, technology. * p < 0.10, ** p < 0.05, *** p < 0.01.

Figure 4.2 depicts the mean review scores on *Kununu* and *Glassdoor* for each industry. With the exception of the energy industry, the average review score is always higher on *Glassdoor* than on *Kununu*. The differences are significant for the automotive (unpaired t-test (unequal variances), *p*-value = 0.0298, N = 16 companies), consumables (unpaired t-test (equal variances), *p*-value = 0.0173, N = 13 companies) and technology industry (unpaired t-test (equal variances), *p*-value = 0.0505, N = 8 companies), and also for all industries combined (unpaired t-test (unequal variances), *p*-value < 0.001, N = 103 companies).

How can these results be explained? In the light of employees' possible concerns regarding anonymity it seems plausible that employees who provide a more negative review of their company on *Glassdoor* are less likely to specify a clear company location compared to

employees who provide a more positive review. Consequently, the results shown in Figure 4.2 can be explained by the fact that the average review scores included for *Glassdoor* are based on a subset of comparatively good reviews, whereas for *Kununu,* the whole range of reviews is included. This explanation is also supported if one compares *Glassdoor's* average review scores for reviews with and without an indicated company location. For reviews in German language with an indicated company location, the mean of the average review score is 3.84 whereas for reviews in German language without an indicated company location, the mean is 3.78. However, a comparison of these values from *Glassdoor* is problematic since the reviews with an indicated company location belong to the set of all reviews in German language.

4.5.2 Self-selection and time of posting a review

The provision of a company review on an ERP is voluntary. The reviews submitted are not random samples of the workforce and are therefore subject to a self-selection bias. Employees only write a review if the benefit they feel from doing so outweighs the effort involved in writing it. For reviews on online marketplaces, various authors (Marinescu et al., 2021; Luca and Zervas, 2016; Masterov et al., 2015; Hu et al., 2009) point out that the majority of reviews is written by users who made a particularly positive or negative product-related experience. Very positive reviews are usually observed more frequently than very negative reviews. People are open to share information on the internet, especially in states of arousal (Berger and Milkman, 2012; Berger, 2011). Therefore, for ERPs, one might assume that at least some employees feel the need to share their work-related experiences and opinions online, especially at times when work-related experiences take on an above-average positive or negative form.

If this assumption is applicable to a fraction of employees, this would imply that employees with moderate work-related experiences are under-represented on ERPs compared to their actual distribution in a company's workforce. However, for *Glassdoor*, Marinescu et al. (2021) find that the distribution of reviews is relatively balanced.

Users of online marketplaces are usually asked to provide a review in a message (e.g., by E-mail or messenger services such as *WhatsApp*) immediately after completing a transaction. The evaluation is thus made at a time when users are likely to remember the transaction relatively well. In contrast, on ERPs it is much more difficult to make a statement about at which point of time in their career, employees decide to rate their company.

Imagine, for example, an employee after a job change who was far less satisfied with her former affiliated company than with her current company. After two months this employee rates her current company on an ERP benevolently and positively. After a further six months, the employee has settled into the environment of the new company and now views her job far less euphorically than in the first months. This process is known as hedonic adaption (or hedonic treadmill) (see e.g., Frederick and Loewenstein, 1999). It describes the phenomenon that after a positive or negative evaluated life change, the satisfaction level of a person will approach its original level after a relatively short time. In the example above, the employee's level of satisfaction, which has now fallen again, would not be reflected in her original review.

Based on the reviews on ERPs, no statement can be made about the degree of self-selection. Further, without a detailed qualitative analysis of single reviews, it is not possible to determine in which emotional state the reviewer was or how long she had been working for the evaluated company when she wrote the review. However, *Kununu's* filter functions allow aggregated review scores to be generated for both current and former employees. In order to get an approximate idea of whether the time of posting a review has a relevant impact on the average reviews, the average review scores of both current and former employees are compared below.

Over all companies, the average review scores for current employees (mean = 3.71, sd = 0.25) are significantly better (unpaired t-test (unequal variances), p-value < 0.001, N = 111 companies) than for former employees (mean = 3.33, sd = 0.43). This result is robust to all industry classifications in the data set (unpaired t-tests, all p-values < 0.1). This result is, however, not entirely surprising since it can be assumed that many former employees have left a company precisely because of dissatisfaction and therefore rate this company worse than current employees.

However, the results illustrate that a comparison of the aggregate ERP reviews of companies can be problematic if the percentage share of former employees' reviews among all reviews differs between these companies. For example, depending on the company, the percentage share of former employees' reviews among all reviews in the automotive industry ranges from 16.29 to 40.04% (mean = 25.52%, sd = 7.12%, N = 16 companies).

4.5.3 Employee's awareness of their impact on a company's reputation

Employees who rate a company on an ERP influence the reputation of that company. The company reviews on ERPs are read both by job-seekers and the current employees of a company. Studies (Wayne and Casper, 2012; Chapman et al., 2005) indicate that a good

reputation increases the attractiveness of a company for job-seekers. A good reputation of a company attracts a higher number of job-seekers compared to companies with a lower reputation (Turban and Cable, 2003). In addition, a higher reputation attracts job-seekers with higher abilities (Bidwell et al., 2015). Current employees are also influenced by the reputation of their affiliated company. Their own engagement is positively influenced by the company's reputation (Shirin and Kleyn, 2017; Men, 2012) and they are more likely to remain in their company through well-established HRM practices (App et al., 2012). Arnold and Staffelbach (2012) show that employees who trust their employer and who have a high level of perceived employability show lower levels of job insecurities after a company restructuring.

When providing a company review on an ERP it is likely that at least a fraction of employees has an interest in maintaining the already good reputation or to increase the reputation of their affiliated company. Since co-workers with high abilities who fit well into the company can help to maintain and further improve a company's good reputation, the current employees of a company may benefit from their company recruiting the best possible applicants for vacant positions. In order to attract applicants, it is advantageous for companies to have a good reputation on an ERP. For these reasons, current employees have strong incentives to influence the reputation of their company in the most positive way. Helm (2011) examines which factors influence employees' awareness of their impact on a company's reputation. Her findings show that especially the pride employees feel for being affiliated with a company has a positive effect.

Considering the entire workforce of a company, it is unclear to what extent individual employees or employees in different positions are aware of their impact on the company's reputation. However, media reports on companies focus particularly often on the management personalities of companies. In addition, various studies (Conte, 2018; Love et al., 2017) examine the influence that executives (especially CEOs) have on the reputation of companies. It is therefore likely that employees in executive positions are particularly aware of their influence on a company's reputation. We would therefore expect the average reviews of executives to be better than those of non-executives.

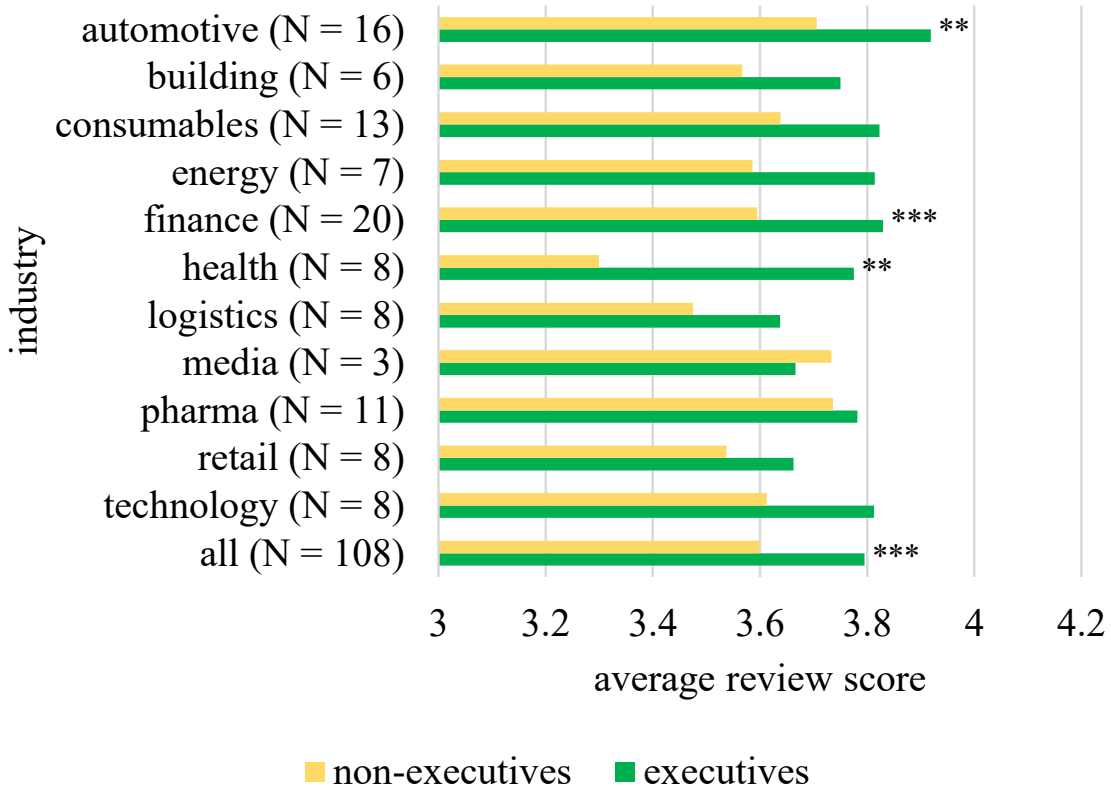


Figure 4.3: Mean review scores of non-executives and executives on Kununu ordered by industry.

Notes. Unpaired t-tests (unequal variances) for the following industries: automotive, building, logistics, all. Unpaired t-tests (equal variances) for the following industries: consumables, energy, finance, health, media, pharma, retail, technology. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 4.3 shows the average review scores for the reviews of non-executives and executives on *Kununu*. The average review scores of executives are better than those of non-executives in each industry, except for the media industry for which only three companies are included in the data set. Significant differences can be observed for the automotive, finance, and health industry as well as for all industries combined. Regarding these results, it must be emphasized that based on the data it cannot be distinguished whether executives provide better average reviews because they have a stronger awareness of their impact on a company's reputation or whether they actually perceive their job as better because of e.g., above-average salaries and/or their prominent position within the company. Nevertheless, these results clearly show that the average reviews of employees in different positions can differ significantly. Consequently, these results further demonstrate that a comparison of the aggregate review scores of companies can be problematic if the percentage share of reviews from executives among all reviews deviates between these companies.

4.5.4 Socially influenced preferences

In their product- or service-related preferences, individuals are often influenced by the existing preferences of other people (Cialdini and Goldstein, 2004). A deviation of one's own preferences from other people's preferences can cause a state of cognitive imbalance. According to balance theory (Heider, 1946), people tend to adjust their attitudes towards the evaluated circumstances or objects, or adjust their attitudes towards others in order to achieve a more balanced state of mind.

For the submission of employer reviews, it seems reasonable to assume that employees do not exclusively consider own work-related opinions, but are influenced by the existing reviews of their current or former colleagues. The extent to which an employee's own opinion is influenced by existing reviews may also depend on the degree of sympathy an employee has with her current or former colleagues. Izuma and Adolphs (2013) have experimentally demonstrated that students improved their original rating of a t-shirt after they were told that their fellow students who were perceived as sympathetic rated the t-shirt better than themselves. At the same time, students downgraded their original rating of a t-shirt after learning that sex offenders who were perceived as unsympathetic had rated t-shirts similarly well. Concerning how long people maintain this change of attitudes, there is conflicting evidence. While Izuma and Adolphs (2013) observed that preferences were still socially influenced after 4 months, Huang et al. (2014) found that such an effect was only noticeable for a few days before the subjects returned to their original preferences.

For the reviews on online marketplaces like *Amazon* or *eBay*, it can be assumed that social conformity pressure is only of extremely minor importance since the individual reviews come from people who are usually not connected to each other in any way. Reviews on ERPs are submitted anonymously, but compared to online marketplaces the social distance is much smaller. It certainly seems plausible that employees, especially based on job characteristics (e.g., position, department) included in single reviews, identify themselves with the persons who wrote existing reviews and therefore unconsciously give a better or worse review than they would have done without this priming.

However, based on the data collected for this paper, no evidence can be obtained as to whether and to what extent socially influenced preferences affect employee reviews on ERPs.

4.5.5 Regression analysis

Section 4.4 showed that for *Kununu* 88 and for *Glassdoor* 79 of the 114 companies in the data set are marked as active employers on the respective ERP. In the following, a regression analysis is used for both ERPs to test whether the average review scores of companies that are marked as active employers differ from the average review scores of companies that are not marked as active employers. For *Kununu*, it is further investigated whether the share of former employees' reviews and the share of executives' reviews still have an influence on the average review scores when including relevant control variables in the regression.

	Kununu			Glassdoor		
	(1)	(2)	(3)	(4)	(5)	(6)
Active employer	0.07	0.09	0.14**	0.03	-0.03	-0.04
	(0.06)	(0.07)	(0.07)	(0.12)	(0.13)	(0.14)
Months since First Review	0.01***	0.01***	0.00*	0.00	0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Reviews from former employees (%)		-0.01*	-0.01***	data not available		
		(0.00)	(0.00)			
Reviews from executives (%)		-0.01	-0.01*	data not available		
		(0.00)	(0.00)			
Employees (in 1000)		0.00	-0.00		-0.00	-0.00
		(0.00)	(0.00)		(0.00)	(0.00)
DAX			0.06			0.20*
			(0.05)			(0.10)
Automotive			0.08			0.12
			(0.09)			(0.21)
Retail			-0.08			-0.04
			(0.16)			(0.24)
Energy			-0.08			-0.33
			(0.13)			(0.25)
Pharma			0.06			-0.04
			(0.11)			(0.24)
Building			-0.11			-0.21
			(0.10)			(0.27)
Technology			-0.08			0.15
			(0.12)			(0.24)
Consumables			-0.01			0.25
			(0.13)			(0.22)
Logistics			-0.20			-0.24
			(0.11)			(0.25)
Finance			-0.13			-0.16
			(0.08)			(0.20)
Health			-0.34**			-0.80**
			(0.15)			(0.54)
Constant	2.79***	3.13***	3.47***	3.64***	3.63***	4.09***
	(0.21)	(0.29)	(0.40)	(0.30)	(0.35)	(0.37)
N	113	97	97	103	89	89
R^2	0.08	0.15	0.33	0.01	0.01	0.27

Table 4.2: OLS regression of average review scores on ERP characteristics, company characteristics, and clustered industries.

Notes. The numbers (1) to (6) refer to different regression models. In all models the dependent variable is the average review score. The independent variables in models (1) and (4) refer directly to ERP characteristics of the companies. Where possible, models (2) and (5) additionally take into account further ERP characteristics as well as the number of employees (in 1000) for each company. Models (3) and (6) further take into account whether a company has been listed in the German stock index DAX since 2007 and whether a company is classified in the respective clustered industry. Robust standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

In the OLS regression models in Table 4.2, the average review score on *Kununu* is the dependent variable in models 1-3 while the average review score on *Glassdoor* represents the dependent variable in models 4-6. The upper four independent variables relate to the ERPs directly. The variable *active employer* is a dummy variable that takes a value of 1 if a company is marked as an active employer on the respective ERP. The variable *months since first review* counts the number of months since the first review on each ERP. *Reviews from former employees (%)* contains the percentage share of reviews from former employees among all

reviews. Likewise, *reviews from executives (%)* contains the percentage share of reviews from executives among all reviews. A further control variable is *employees (in 1000)* which is an average of the average number of employees in the years 2016 and 2014. The remaining control variables refer to company characteristics and are dummy variables taking a value of 1 if a company has been listed in the German stock index *DAX* since 2007 or is classified in the respective clustered industry.

For *Kununu*, no robust effect of an active profile can be observed. Only in model 3, an active profile positively affects (p-value < 0.05) the average review score of a company. For *Glassdoor,* no effect of an active profile can be observed in any model. It has to be noted that there is no data available on how long a company has had an active employer profile. Therefore, based on the results from Table 4.2, no conclusive statement can be made as to whether an active employer profile has an effect on average review scores. The variable *months since first review* has a positive effect for *Kununu* in all models, but not in any of the models for *Glassdoor*. For *Kununu*, a significantly negative effect can be observed for the percentage share of reviews from former employees. This result implies that a 10% increase in the percentage share of reviews from former employees on average leads to a reduction of the average review score by 0.1 stars. The variable *reviews from executives (%)* has a significantly negative effect in model 3 but not in model 2. This result differs from the result in Figure 4.3, where the average review scores of executives were significantly better than the average review scores of non-executives. However, the negative coefficient of *reviews from executives (%)* is significant only at the 10% level and the regression in Table 4.2 contains a number of control variables that are not included in Figure 4.3. Therefore, we refrain from a further interpretation of this result. Additionally, for *Kununu* and *Glassdoor*, negative effects (p-value < 0.05) for the industry *health* can be observed.

4.6 Suggestions for informative ERP designs

4.6.1 How to address possible bias factors

The results from the previous section have shown that the perceived level of anonymity can affect a review's informativeness. It was highlighted that the concrete design of an ERP has an influence on the average review scores. It seems reasonable that *Glassdoor* aims to motivate as many employees as possible to provide a review by offering a voluntary option to indicate a company's location. Especially for international companies, it is often unclear to ERP users to which location or geographical area a review refers. In the case of small companies, the

possibility to deliberately avoid specifying the company's location appears reasonable as it guarantees the anonymity of the reviewer. However, for the reviews of larger companies, it is quite questionable what benefit individual reviews have if it is unclear whether these reviews relate to locations in e.g., Portugal, Brazil or Germany.

ERP operators have to weigh up how they can persuade employees to provide informative reviews whilst accounting for their concerns regarding anonymity. To address this issue, ERPs could oblige employees to indicate a company location when providing the review, but give them an option not to publish the company location publicly with their review. In this way, the review scores and further related values of such a review could at least be included in the aggregated scores of a specific company location.

Employees from smaller companies might fear that they could easily be identified by their bosses or colleagues through their ERP review. To address such concerns, ERPs could give these employees the option to only include their review in the aggregated score of their company and not as a separately visible review. It is further conceivable that such reviews could be disclosed as visible reviews only after at least a certain number of reviews have been provided for the respective company since an individual assignment to reviews would then be less likely.

The next suggestions refer to self-selection and the time of posting a review. Self-selection of particular groups of employees could mainly be reduced if companies would actively encourage their whole workforce to provide reviews on ERPs. ERPs could focus their marketing efforts particularly on those groups of employees who are currently underrepresented on the respective ERP. Regarding the time of posting a review, it has been shown that the reviews of current and former employees differ. In order to allow a better comparison between companies, ERPs could set their filter defaults in such a way that initially only the average review scores of current employees are displayed. The possibility that a fraction of reviews have been written in aroused states could be reduced by asking employees of ERPs to verify or renew their submitted reviews regularly. Reviews that are regularly verified or updated by the same employee could be flagged as highly informative by ERPs. By re-examining her first review, an employee might register if she had written her first review with too much euphoria or anger and accordingly correct the first review if necessary. The ERP could then calculate an average review score from the individual reviews of an employee in order to prevent the reviews from counting more than once.

A useful suggestion regarding the possible biasing impact of an employees' awareness of her impact on a company's reputation is more complicated. Without the use of time-consuming

questionnaires (e.g., in Helm, 2011) it is impossible to determine an employees' awareness of her impact on a company's reputation and even with a detailed questionnaire, a socially desirable response behavior cannot be ruled out. However, on the profile pages of individual companies, ERPs could highlight which percentage shares of the reviews were provided by which groups of employees. ERPs could further indicate when average review scores differ particularly strong between different groups of employees (or between different locations of the same company).

An additional suggestion refers to the socially influenced preferences of employees. ERPs could increase the informativeness of reviews by trying to prevent a possible priming through already existing reviews. *Glassdoors'* "Give to get" policy partly helps to reduce a possible priming effect as employees have only limited access to contents when they visit the *Glassdoor* website for the first time. Nevertheless, it would make sense for ERPs to consider a design in which users when opening a company's ERP site are asked whether they just want to inform themselves or if they want to rate their company first. In the second case, users could be reminded that in order to capture unbiased opinions, it would be useful for them to write a review before reading any of the other existing reviews.

4.6.2 Further suggestions

To ensure the best possible matching process between employees and companies via ERPs, it would be beneficial to employees if they could individually weight which attributes are particularly important to them in the search process for a suitable company. Suitable companies could then be presented in a ranking based on the individual weightings and the already submitted reviews.

A company's average review score displayed by ERPs is based on all reviews submitted since the existence of a company's site on the ERP. If a user now compares the average review scores of different companies on an ERP, the average review scores are the result of reviews that have not been written within the same time frame. By using the filter function of *Kununu*, it is possible to display a company's average review score of the past month, the past 6 months, and the past 12 months. Here, it is suggested that ERPs could display the average review score of the last 12 or 24 months by default. Job-seekers would benefit from this by being able to compare companies' current working conditions with each other without having to set a filter first. Such a design feature would also strengthen the incentive for companies to improve their employer quality. First, companies could not rely on good reviews older than 12 or 24 months. Second, implemented quality improvements by companies would also become visible more

quickly since poor reviews older than 12 or 24 months would no longer be included in the average review score.

Especially for larger companies where employees may not have to be very concerned about maintaining their anonymity, it would be reasonable to specify demographic variables when submitting a review on an ERP. Drabe et al. (2015) show that job satisfaction varies between different age groups and that different age groups attach importance to different factors with regard to their job satisfaction. Therefore it would be useful if, for example, an older employee could specify on an ERP via a filter that she only wants to see reviews of employees older than 45 years.

ERPs could additionally enable registered users to mark reviews from other users as helpful. In this way, ERP users who have posted a review would be informed whether their review was perceived as helpful which could motivate them to provide more (informative) reviews in the future.

Further, it would be useful if current and former employees could voluntarily state in their reviews how long they have been working or have worked for the rated company and how many companies they have worked for previously. Based on this information, ERPs could present details about how long former employees have worked for that company on average and thereby provide ERP users with an indication of a company's employee turnover rate. Additionally, former employees could be asked on a voluntary basis about the reasons why they left a company.

4.7 Conclusion, limitations, and suggested research agenda

This article examined specific design features of ERPs in detail. By consulting the relevant literature, it was shown that the rating environment of ERPs differs substantially from the well-studied rating environments of online marketplaces. Possible bias factors such as the perceived level of anonymity and the timing of review provision resulting from the special rating environment of ERPs were discussed. Whenever possible, it was empirically demonstrated that these factors can have an influence on aggregated review scores. Suggestions on how to address the problems connected with these bias factors were presented. Additionally, further suggestions for more informative ERP designs were outlined.

This paper has a number of limitations. At the level of the individual reviews, there was no control on when these reviews were provided. In particular, the results from the comparison of

the average review scores of *Kununu* and *Glassdoor* should be treated with caution as *Kununu* has been active in Germany for a much longer time. To perform the regression analysis with as many reviews as possible, the average review scores based on the complete review period were chosen as the dependent variable. The dependent variable was recorded at a fixed date in April 2020, whereas the independent variable *employees (in 1000)* refers to dates several years earlier. In addition, whether the companies were marked as active employers on the respective ERP was also recorded on a fixed date. Therefore, no statement can be made as to when this activity started or whether inactive companies were active on the respective ERP in the past. The data set contains ERP values and company key figures for 114 large companies that are active in Germany. Therefore, it is unclear whether the differences between the average review scores on *Kununu* and *Glassdoor* and between the different subgroups on *Kununu* also apply to smaller companies and/or companies outside Germany.

ERPs and the information provided on them offer numerous perspectives for future research. Similar to the study by Marinescu et al. (2021), the effects of implemented design changes on ERPs could be examined more closely. Since the (attempted) posting of counterfeit reviews can be observed on many platforms (Luca and Zervas, 2016; Mayzlin et al., 2014), related to ERPs this issue also presents a promising field of research. Since laboratory experiments allow to control for a wide range of confounds (see e.g., Cloos et al., 2020; Weimann and Brosig-Koch, 2019), they could be used to investigate in detail the extent to which factors such as socially influenced preferences influence the evaluation of one's own employer.

The information provided on ERPs could also be used to extend existing research on corporate social responsibility (see e.g., Fietze et al., 2019; Henry and Möllering, 2019; Uzhegova et al., 2019). At the level of individual companies, future studies could examine whether the existing level of corporate social responsibility, or whether and how newly implemented corporate social responsibility initiatives have an effect on a company’s reviews. Furthermore, the question of how platform users interpret the ratings of ERPs compared to the ratings on online marketplaces would offer an important and interesting field of research which has not been investigated yet.

Acknowledgements

I thank the guest editor Sven Hauff and two anonymous referees for very helpful feedback. In addition I thank Matthias Greiff for his valuable comments and for encouraging me to complete this article. Remarks of Merindah Loessl and Svenja Mohr improved the article considerably.

4.8 Appendix

Industry Classification of the Companies in the Data Set

Automotive / Supplier / Mechanical Engineering (N = 16): Volkswagen AG, Daimler AG, BMW AG, Robert Bosch GmbH, INA-Holding Schaeffler GmbH & Co. KG, ZF Friedrichshafen AG, Ford-Group Germany, Opel Automobile GmbH, MAHLE GmbH, Freudenberg & Co. KG, Continental AG, Liebherr-International-Gruppe Deutschland, ABB-Gruppe Deutschland, Hella KGaA Hueck & Co., Krones AG, Voith GmbH

Retail (N = 9): Rewe-Gruppe, Edeka-AG, Aldi-Süd, METRO AG, Otto Group, dm-drogerie markt, Tchibo, Lidl, Kaufland

Energy / Water- / Waste Management (N = 8): EWE AG, Stadtwerke München GmbH, RWE AG, E.ON SE, Vattenfall-Gruppe Deutschland, EnBW Energie Baden-Württemberg AG, Rethmann SE & Co. KG (Remondis), Stadtwerke Köln GmbH

Pharma / Chemistry (N = 13): Bayer AG, BASF SE, Fresenius SE & Co. KGaA, Evonik Industries AG, C. H. Boehringer Sohn AG & Co. KG, MERCK KGaA, LANXESS AG, Roche-Gruppe Deutschland, Sanofi-Gruppe Deutschland, Wacker Chemie AG, B. Braun Melsungen AG, Lyondellbasell-Gruppe Deutschland, Bilfinger SE

Building- / Raw Materials / Steel (N = 7): Adolf Würth GmbH & Co. KG, thyssenkrupp AG, Salzgitter AG, STRABAG Gruppe Deutschland, Saint-Gobain-Gruppe Deutschland, VINCI-Gruppe Deutschland, K+S AG

Technology / Telecommunications (N = 8): Linde AG, SAP SE, Siemens AG, Deutsche Telekom AG, IBM-Gruppe Deutschland, Carl Zeiss AG, United Internet AG (1&1), HP-Gruppe Deutschland

Consumables (N = 14): BP-Gruppe Deutschland, Henkel AG & Co. KGaA, Dr. August Oetker KG, Shell-Gruppe Deutschland, BSH Hausgeräte GmbH, Tchibo GmbH, Beiersdorf AG, Procter & Gamble-Gruppe Deutschland, Nestlé-Gruppe Deutschland, INGKA-Gruppe Deutschland (IKEA), Miele & Cie. KG, Philip Morris International-Gruppe Deutschland, TOTAL-Gruppe Deutschland, H & M Hennes & Mauritz-Gruppe Deutschland

Health / Other Services (N = 10): Asklepios Kliniken GmbH, Sana Kliniken AG, Adecco-Gruppe Deutschland, AVECO Holding AG (WISAG), DEKRA SE, Vivantes - Netzwerk für Gesundheit GmbH, DFS Deutsche Flugsicherung GmbH, Rhön-Klinikum AG, Charité Universitätsmedizin Berlin KöR, Kühne + Nagel-Gruppe Deutschland

Logistics / Defence / Transportation (N = 7): Deutsche Bahn AG, Deutsche Post DHL, Deutsche Lufthansa AG, Airbus-Gruppe Deutschland, Fraport AG, Rheinmetall AG, Rolls-Royce-Gruppe Deutschland,

Finance / Consulting / Insurance / Investment (N = 21): Commerzbank AG, Allianz SE, Deutsche Bank AG, Münchener Rückversicherungsgesellschaft AG, Deutsche Börse AG, KPMG AG, Norddeutsche Landesbank Girozentrale, Landesbank Baden-Württemberg, Bayerische Landesbank, HDI Haftpflichtverband der Deutschen Industrie V.a.G., Ernst & Young-Gruppe Deutschland, KfW Bankengruppe, HGV Hamburger Gesellschaft für Vermögens- und Beteiligungsmanagement mbH, UniCredit-Gruppe Deutschland (HypoVereinsbank), DZ Bank AG, PricewaterhouseCoopers AG, AXA-Gruppe Deutschland, HUK-COBURG, Debeka-Gruppe, Signal-Iduna Gruppe, Vonovia SE

Media (N = 3): Bertelsmann SE & Co. KGaA, Axel Springer SE, ProSiebenSat.1 Media SE

5. Acceptance of data sharing in smartphone apps from key industries of the digital transformation: A representative population survey for Germany

This chapter has appeared as:

Mohr, S. & Cloos, J. (2020). Acceptance of data sharing in smartphone apps from key industries of the digital transformation: A representative population survey for Germany. MAGKS Joint Discussion Paper Series in Economics, No. 02-2021.

This paper can be downloaded via

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3779905

Abstract: The use of smartphone apps has numerous advantages for app providers and users. However, the users of many smartphone apps are confronted with a trade-off between usage benefits and preferences for personal data protection. We investigate the acceptability of data sharing in different hypothetical scenarios describing five types of these apps from key industries of the digital transformation. In a representative survey for the German population (N = 1,013), we examine to what extent the acceptance of data sharing is influenced by potential recipients, collected information attributes, and the promoted benefits of data sharing. We differentiate the promoted benefits in two treatments according to monetary (or personal) and environmental (or public) benefits. Our results show no treatment effects but significant differences in acceptance values for different recipients and information attributes. We further observe that participants with stronger green consumption values, participants with a stronger risk propensity, men, and younger participants show a higher acceptance towards data sharing in the described scenarios. (JEL: O33, Q18, C83, L86, M31, M37)

Keywords: privacy, digitalization, digital transformation, representative survey, data protection, environmental attitudes

5.1 Introduction

In the course of the Covid-19 pandemic, governments in various countries have developed digital contract tracing apps. The aim of these apps has been to inform people about possible risk contacts and thus help to slow down the spread of the pandemic. In Germany, large parts of the society have had a positive attitude towards the app. At the same time, intensive debates on various data protection concerns arose even before the apps were launched (Altmann et al., 2020; Amann et al., 2020; Rowe, 2020). The example of the Covid-19 tracing app illustrates important privacy relevant characteristics of apps and at the same time reveals problems that app providers face on a regular basis. Many apps require consumers to disclose personal information to actors of the private or public sector in order to use the apps and the benefits they offer. For the data-providing consumers, there is consequently a trade-off between preferences for their own data protection and the potential benefits of data disclosure (Acquisti et al., 2020, 2016, 2015; Kokolakis, 2017). For app providers, it is essential to know which features of an app influence users' privacy preferences to what extent. Without such knowledge, users' privacy concerns may not be adequately addressed, and users may be skeptical about using an app or could decide against using an app at all. Since individuals' online privacy preferences are highly heterogeneous (Acquisti et al., 2020, 2016, 2015) and context specific (John et al., 2011) knowledge on specific acceptance drivers is crucial.

Acquisti et al. (2016) describe the disclosure of personal data and the protection of this data as two sides of the same coin. The disclosure of personal data is associated with benefits for the consumers who provide the data, e.g., in the form of financial savings when purchasing products or from bonus payments. By using apps, consumers can reduce search costs during shopping and adapt their consumption behavior in order to match their own preferences. Companies benefit because they can increase their profits by collecting information from their consumers. In this way, companies can efficiently utilize existing consumer potential and save resources, for instance, by avoiding excessive advertising. However, for consumers, the disclosure of personal data can also lead to negative consequences, such as identity theft (Moore et al., 2009), discrimination of various kinds (Cui et al., 2020; Edelman et al., 2017), or a burden through excessive advertising (Johnson, 2013). Companies can also suffer from disadvantages, such as costs resulting from data theft (Hinz et al., 2015), or misuse of data by members of their own company or by members of affiliated companies (Acquisti et al., 2016).

Since the early 2000s, various scientific studies have investigated which factors influence privacy preferences in different contexts (for overviews of these studies see

Acquisti et al., 2020, 2016, 2015; Kokolakis, 2017). The results of these studies show, for instance, that the willingness to share data is influenced by the number of potential recipients (Schudy and Utikal, 2017), the content of the data collected (Cloos et al., 2019; Benndorf and Normann, 2018; Schudy and Utikal, 2017), and the survey framing (John et al., 2011). In the context of digital transformation, Apthorpe et al. (2018) use an advanced survey method to examine privacy norms in various smart home settings. The survey method builds on the theory of Contextual Integrity (Nissenbaum, 2009) which states that data protection standards are context-specific and face the generally accepted adequacy of a specific information exchange. Apthorpe et al. (2018) divide the contexts into the parameters sender, receiver, information attribute, transmission attribute or benefit associated with data transmission, and subject, which enables them to combine different information flows. In this way, the authors identify that in certain smart home contexts even the change of a single parameter can have a significant impact on a data protection standard. For example, participants indicate, on average, considerably higher acceptance values for a fitness tracker sharing data on the heart rate of its owner than sharing data on the eating habits of its owner.

Although knowledge about the acceptance of data sharing in specific digital technologies is important for companies, regulatory authorities, and research, this topic has rarely been investigated in real life settings. Laboratory or field experiments on privacy usually focus on single contexts or on completely artificial situations (see Kokolakis, 2017). While the survey investigation by Apthorpe et al. (2018) explicitly refers to the smart home context, little is known about factors that influence the acceptance of data sharing in the industries most affected by digital transformation. For example, previous research has not addressed the question whether specific external benefits of apps lead to higher acceptance for data sharing. Furthermore, there is comparatively little evidence on how acceptance of data sharing is shaped by various socio-demographic factors and personal attitudes.

This paper investigates the acceptance of data sharing in apps for five key industries of digital transformation. The selected industries are retail, health, nutrition, mobility, and energy. With a representative survey (N = 1,013) for the German population we examine in hypothetical but realistic scenarios, how the acceptance of data sharing via apps varies depending on potential recipients (e.g., market research companies, employer, or federal ministries) and information attributes (e.g., live location, nutrition style, or monthly net income). Wright et al. (2014, p. 325) point out that "…scenarios are a useful instrument to provoke policy-makers and other stakeholders, to including industry, in considering the privacy, ethical, social and other

implications of new and emerging technologies." For each scenario within the survey, we first give a brief and concise explanation of what the app does, who the app provider is, what information attributes must be mandatorily provided in order to use the app, and what benefits the app offers. Then, participants have to assess for further optional data recipients how acceptable it is that the information collected with the app will be shared with these recipients. In the final step, participants have to evaluate for various optional information attributes how acceptable it is that these attributes are collected via the app.

The extent to which green advertising strategies and green consumption values affect the privacy preferences of consumers has not been investigated in the literature so far. We therefore collect acceptance values in two different treatments. The treatments differ according to whether the transmission of data is primarily highlighted by monetary (or personal) or by environmental (or public) benefits. The additional collection of demographic variables and personal attitudes to privacy, sustainable consumption, and risk allows us to analyze how these factors affect the acceptance of data sharing. Our method is adapted from Apthorpe et al. (2018). However, with the decisive difference that our study covers five different industries, and in this way demonstrates how the method developed by Apthorpe et al. (2018) can be applied in a wide range of other contexts.

Our results show no treatment effects for recipients and information attributes. In all scenarios, average acceptance values for data sharing differ significantly between different recipients and information attributes. Acceptance values are particularly low for recipients and information attributes that have a low thematic fit with the respective scenario and where data sharing can potentially lead to very negative consequences. We further observe that the acceptance towards data sharing is lower for stronger online privacy preferences and higher for a larger risk propensity and for stronger green consumption values.

The findings of our paper can help app providers from various industries to identify and address sensitive privacy areas and thereby successfully realize the potential of their existing and planned apps. Our paper additionally provides important insights for public authorities and consumer protection agencies that can be used to adequately address consumers' privacy protection issues. Our results also provide impulses for further scientific research in the fields of privacy and (managing) digital transformation.

5.2 Method

5.2.1 Scenario development

As mentioned above, Apthorpe et al. (2018) identify various Internet of Things (IoT) applications in the context of smart homes as a relevant area for research on privacy standards. Acquisti et al. (2016) find that online advertising, price discrimination in different industries, health care, and finance (lending) are relevant areas where a trade-off between benefits through data provision and privacy preferences exists. Online dating platforms, sharing services such as *Airbnb,* and recruitment processes are further mentioned as relevant areas.

The hypothetical scenarios in this study refer to industries that are largely affected by the digital transformation. We selected the industries according to the following criteria: First, each industry should have a connection to daily consumption, shopping, or health behavior. In this way, we intended to ensure that the scenarios described did not appear too abstract to our participants and that the majority of participants were at least partially familiar with the content of the scenarios. Second, in each industry, (tracking) apps should already exist or at least be conceivable. Finally, these apps should bring benefits to customers, but also require sharing personal data. Related literature underlines the increasing importance of smart technologies (e.g., smartphone apps) in our selected industries - retail (e.g., Roy et al., 2017), health (e.g., Tresp et al., 2016), nutrition (e.g., O'Sullivan et al., 2018), mobility (e.g., Del Vecchio et al., 2019), and energy (e.g., Horne et al., 2015). The relevance of the selected industries is further highlighted by the fact that the German digital association Bitkom identifies retail, health, mobility, and energy as key industries of digital transformation (Bitkom, 2020)[31].

In a review paper on scenario planning, Amer et al. (2013) identify internal consistency, plausibility, creativity, and relevance as the most important validation criteria. The scenario development in our study is based on these criteria. In the descriptions of our scenarios, we use a logical and coherent structure in order to achieve a high degree of internal consistency (see Appendix 5.6.1). As we explain in section 5.2.1.1, the apps and technologies in our scenarios are derived from existing apps and technologies. To guarantee a high level of plausibility, we made sure that the information flows we describe are conceivable. Relevance is ensured by basing our study design on the current literature on smart technologies, digital transformation,

[31] Bitkom (2020) further includes the agricultural industry. In our study, we integrate the food industry to ensure consumer orientation.

data analytics, and privacy. Creativity is achieved by describing apps with different functionalities from different industries.

Table 5.1 presents the parameters used in our scenarios. A detailed description of each scenario can be found in Appendix A (Tables 5.5 to 5.9). In the following, we provide a rationale for the selected sending device and provider combinations, recipients, information attributes, and transmission benefits.

Scenario and Provider		Sending device	Information attributes (mandatory)	Recipients (optional)	Information attributes (optional)	Transmission benefit
A	Super-market	Loyalty card with app	- Name - Address - Date of Birth	- Household members^{A,B,C,D,E} - Employer^{A,B,C,D,E} - Federal ministries^{A,B,D,E} - Market research companies^{A,B,C} - German food producers^{A,C} - Local and long-distance public transport companies^{D,E} - City or municipality^{D,E} - Health insurance company^{A} - American food producers^{A} - Chinese food producers^{A} - German sports equipment producers^{B} - German Society for Nutrition^{C} - Neighbors^{E}	- Live location^{A,B,C,E} - Monthly net income^{C,D,E} - Nutrition style^{A,B} - Body weight and height^{A,B} - Food intolerances and allergies^{A,C} - Memberships in a sports club or gym^{B,C} - Profession^{D,E} - Number of steps taken^{A} - Date of birth^{D} - Driving behavior^{D} - Use of power sources^{E}	Product recommendations (T1), Waste avoidance (T2)
B	Health insurance company	Tracking bracelet with app	- Name - Date of birth - Gender			Bonus payment (T1), Tree sponsorship (T2)
C	Federal ministry of health	Nutrition app	- Name - Date of birth - Gender - Body weight and height - Nutritional style			Nutritional recommendations: Health-promoting (T1), Env. friendly (T2)
D	Techno-logy start-up company	Mobility-tracking app	- Name - Gender - Live location - Type of vehicle			Mobility recommendations: Cost- and time saving (T1), Env. friendly (T2)
E	Energy provider	Smart meter with app	- Name - Address - Date of birth			Power usage recommendations: Cost reducing (T1), Env. friendly (T2)

Table 5.1: Description of the individual scenarios in terms of provider, sending device, optional recipients, mandatory and optional information attributes, and transmission benefits.

Notes. The superscript letters at recipients (optional) and information attributes (optional) indicate in which scenarios (A to E) the respective recipients or information attributes are included.

5.2.1.1 Sending device and provider

Based on the above-mentioned literature and further non-scientific reports, we selected five sending device and provider combinations. The criteria for each of these combinations were that they are (1) realistic and relevant for the respective industry and (2) that the devices include a possibility of being connected to an app. In the first sentence of each scenario we explained that the respective app is cost-free. For the choice of providers and recipients, we deliberately selected private companies, public companies, and governmental actors. In line with

Apthorpe et al. (2018), we did not mention specific device names in order to avoid associations with existing devices. Our five sending device and provider combinations are:

A. **Loyalty card with app from supermarket chains:** In the retail industry, single companies or coalitions of companies offer loyalty cards in combination with apps and thereby collect information about the product selection and purchasing behavior of their customers (Wang et al., 2018). Based on this information, customers can receive individualized product recommendations or discounts via these apps (Cortiñas et al., 2008). Both the timing and the topic can be specifically targeted towards customers in order to achieve the greatest possible effectiveness of the product recommendations and advertisements (Fernández-Rovira et al., 2021; Acquisti et al., 2016). Smart retail technologies can improve customers' shopping experience (Minch, 2015), e.g., through personalization, but at the same time they also raise privacy concerns (Roy et al., 2017). For companies, like supermarket chains, customer-oriented technologies are an essential tool to attract new customers and to stimulate the purchasing behavior of existing customers (Inman and Nikolova, 2017). Loyalty cards are widely used and well-known in German retail. Our retail scenario is very similar to the loyalty program Payback[32]. In Germany, more than 31 million people currently use the Payback card and 10 million of them actively use the Payback app (Payback, 2020). In our retail scenario, the app providers (supermarket chains) are from the private sector.
B. **Tracking bracelet with app from a health insurance company:** Digitalization affects the healthcare system in various ways (see e.g., Tresp et al., 2016; Agarwal et al., 2010). Fitness tracking apps are not only used to improve the quality and cost of healthcare (Mehta and Pandit, 2018), but are also applied to self-track sport activities and one's personal health (Williamson, 2015). However, in addition to these benefits, sharing personal fitness data and health data can also raise privacy concerns (Vitak et al., 2018) or lead to discrimination among minorities (Joy et al., 2020). German health insurance companies already use data from app-based activity trackers and provide premiums based on this data (Techniker Krankenkasse, 2020). In a representative study for Germany, the market research company Splendid Research (2019) found that 33% (23%) of the German population uses (is interested in) apps or wearables to track personal fitness, health, or nutrition data. 38% of the respondents totally reject the use of these self-measurement

[32] Payback is a multinational and multi-industry bonus system with a customer card and the leading bonus program in Germany (Payback, 2020).

systems. The results of the survey further showed that more than half of the participants would share health-related data with health insurance companies in order to obtain discounts. In our scenario, we did not specify whether the health insurance company is a private or statutory health insurance, since in the German multi-payer healthcare system more than 69 million people (> 83% of the German population) are insured by statutory health insurance companies (Bundesministerium für Gesundheit, 2019: 109).

C. **Nutrition app from the Federal Ministry of Health:** As consumer behavior changes towards self-optimization, there is also a demand for food products that are tailored to individual needs (O'Sullivan et al., 2018; Poutanen et al., 2017). Artificial intelligence and smartphone apps enable practical and personalized nutritional recommendations based on genetic and behavioral information such as eating behavior and physical activity. These nutritional recommendations can help, for instance, to prevent obesity or diseases such as diabetes (Chatelan et al., 2019). In the field of nutrition, there are various apps that enable users to count calories, track their purchased food using barcode scanners, or create personalized nutrition plans. Often these apps can also be combined with other fitness apps (for an overview see DiFilippo et al., 2015). In our nutrition scenario, the app provider (Federal Ministry of Health) is a governmental institution.

D. **Mobility tracking app from a technology start-up company:** Smartphone based tracking in the mobility sector enables to improve urban planning and transport systems and to effectively satisfy people's travel needs (Longhi and Nanni, 2020; Wahlström et al., 2017; Gisdakis et al., 2014; Guido et al., 2012, Mohan et al., 2008). Behavioral changes can be induced by providing consumers with personalized transport solutions. However, mobility tracking requires consumers to disclose sensitive data, such as their live location, which potentially entails privacy concerns (Bucher et al., 2019; Cellina et al., 2019; Del Vecchio et al., 2019; Iqbal and Lim, 2010). Mobility tracking is comparatively less popular in Germany than in other countries. The tracking of car driving behavior to determine user-dependent insurance rates, known as telematics (Longhi and Nanni, 2020; Wahlström et al., 2017), is estimated to be used by less than 1% of all car drivers in Germany.[33] In 2016, by contrast, 17% of Italian, 10% of South African, and 6% of US car

[33] See (in German) https://www.capital.de/geld-versicherungen/telematik-tarife-der-versicherer-faehrt-mit, accessed December 01, 2020.

drivers had already signed up to telematics-based insurance policies.[34] In our mobility scenario, the app provider is a technology start-up from the private sector.

E. Smart meter with app from an energy provider: A smart meter is an intelligent digital electricity meter that records and stores data on power consumption at any time and can also send the stored data (Zheng et al., 2013). Since energy consumption data is automatically and frequently transmitted to the energy provider, smart meters have the potential to raise privacy concerns (Horne et al., 2015). Greveler et al. (2012) show that high-resolution data on a household's energy consumption enables undesired identification and monitoring of the appliances used in the consumer's home. Since 2020, Germany has an obligation to install smart meters if annual electricity consumption exceeds 6,000 kWh. From the year 2032 onwards, smart meters will be mandatory for all households (Bundesministerium für Wirtschaft und Energie, 2020). Energy providers in Germany are often publicly owned companies. However, due to space reasons we refrained from a more specific description of the company in our scenario.

5.2.1.2 Recipients and information attributes

For our scenarios, we selected a broad range of different organizations and groups as potential recipients of the collected information. In each scenario, the participants were asked to what extent they find it acceptable that the information collected by the sending device is shared with optional recipients, in addition to the app provider. After that, the participants were asked to what extent they find it acceptable that the app provider requests and collects optional information attributes with the app in addition to the set of information attributes which has to be mandatorily provided when using the app. In each scenario, we described that possible data sharing with optional recipients and the possible request and collection of optional information attributes was clearly stated in the app's general terms and conditions.

We aimed to provide a good balance for the selected recipients and information attributes and therefore always selected recipients and information attributes for which we expected comparatively high and comparatively low acceptance values. In general, we chose recipients and information attributes that, at least in a broad sense, thematically fitted the respective scenario. We aimed to avoid that the scenarios appeared too unrealistic to the participants since this might have resulted in high dropout rates. For this reason, some optional recipients and

[34] See https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/telematics-poised-for-strong-global-growth, accessed December 01, 2020.

information attributes were only included in one or two scenarios while others were included in each scenario.

5.2.1.3 Transmission benefits

As outlined above, the provision of apps through companies and other organizations can help them to use resources more efficiently and at the same time to purposefully address the needs of their customers or target groups. For users, apps similar to those in our scenarios often provide information and offers that can result in cost and time savings. In each of our scenarios, we described specific benefits that result from data sharing in the respective app. In treatment 1 (T1) monetary (personal) benefits, such as bonus payments or personalized cost-saving behavior recommendations for the app user, were mentioned. In contrast, treatment 2 (T2) mentioned environmental (public) benefits, such as nature-friendly activities through the app provider, that increase when the app is used by a larger number of people or personalized environmentally friendly behavior recommendations for the app user.

5.2.2 Survey design

Our survey consisted of four stages (Figure 5.1). In stage 1, socio-demographic questions on gender, age, residence (federal state), and education were asked in order to verify the quotas of the representative survey.[35] Then, the participants were exposed to five hypothetical scenarios in stage 2. Before presenting the scenarios, participants were randomly allocated into two treatment groups.

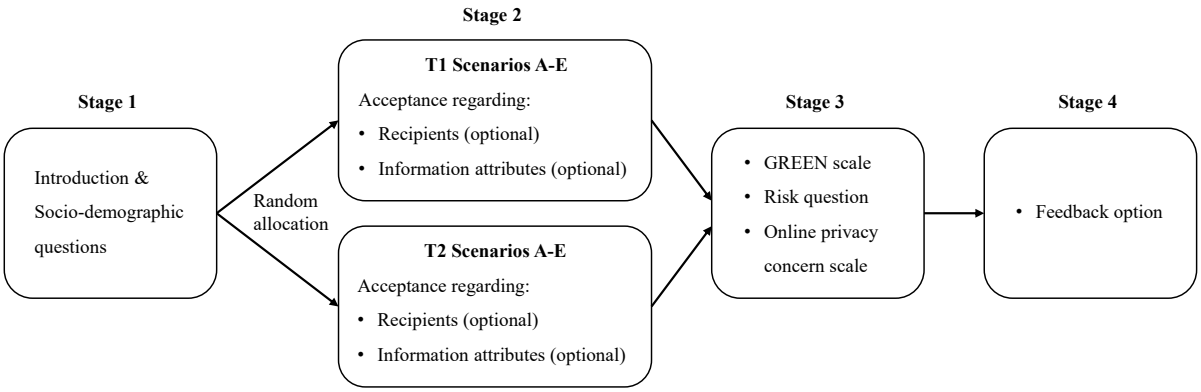


Figure 5.1: Sequence of the survey.

Participants completed all scenarios in either T1 or T2. We used this between subjects design in order to prevent participants from anticipating our research objective. If participants had noticed that we wanted to investigate the effects of monetary and environmental benefits on the

[35] The quotas were calculated based on the database (Genesis online) of the Federal Statistical Office of Germany.

acceptance of data sharing, these participants could have adjusted their response behavior accordingly which could bias our results. The sequence of the scenario presentation corresponded to the described conceptualization in section 5.2.1. In each scenario, participants had to assess the acceptance of data sharing with (1) optional recipients and (2) for optional information attributes. The different recipients and information attributes were presented in a random order.

In total, each participant had to assess the acceptance of 30 recipients and 21 information attributes in 10 boxes. In each scenario, the first box contained different recipients, followed by a second box with information attributes. This order should ensure that participants only considered mandatory information attributes, rather than optional ones, when assigning acceptance values to recipients. Acceptance was measured on a six-level scale from (1) completely unacceptable, (2) somewhat unacceptable, (3) rather unacceptable, (4) rather acceptable, (5) somewhat acceptable to (6) completely acceptable. In order to illustrate our results in section 5.3 as clearly as possible, we converted the original values of our six-level acceptance scale into a range from -1 to 1. Therefore, the negative range includes the values (-1) completely unacceptable, (-0.6) somewhat unacceptable, and (-0.2) rather unacceptable, whereas the positive range includes the values (0.2) rather acceptable, (0.6) somewhat acceptable, and (1) completely acceptable. We deliberately refrained from providing a central answer option in order to avoid an anchor effect towards the middle and to force participants to make explicit acceptance decisions.

In stage 3, participants' attitudes were assessed by using well-validated measures.[36] To collect data on participants' sustainable consumption attitudes, we used the six-item GREEN consumption scale (Haws et al., 2014). We included a quality check question to this scale in order to expose those respondents who did not seriously answer our survey. Participants' online privacy concern was assessed by using a six-item scale by Ham (2017) adapted from Dolnicar and Jordaan (2007). Both latent constructs were measured on seven-point Likert scales ranging from (1) strongly disagree to (7) strongly agree. Lastly, we measured participants' risk attitude by using a single question proposed by Dohmen et al. (2011). Stage 4 closed the survey with a field for optional comments.

[36] Details of both measurement scales are described in appendix D (section 5.6.4).

5.2.3 Hypotheses

5.2.3.1 Treatment effects

Our first set of hypotheses focuses on possible treatment effects. Over the past couple of years, it has been observed that companies frequently underline their environmental or social commitment within their product advertisements. One prominent example is the rainforest project of the German *Krombacher* brewery. In cooperation with the *World Wildlife Fund* (WWF), the brewery promotes its products by saying that customers save one square meter of rainforest by buying a box of beer (Mögele and Tropp, 2010). In a field experiment, Asensio and Delmas (2016, 2015) observe that participants adopt a more environmentally friendly energy consumption behavior if they receive information on the environmental and health impact of their households' energy consumption behavior. In contrast, only small or no changes in consumption behavior can be found when participants receive information on the financial benefits of a more efficient energy consumption behavior. In our study, we expect participants to view data sharing as more acceptable if the environmental benefits of apps are highlighted. We further expect acceptance values to be higher in T2, as participants may consider data sharing in T2 more acceptable due to image concerns (Ariely et al., 2009). In T2, participants could tend to evaluate the apps described in the scenarios as more acceptable to express their prosocial and environmentally friendly attitude towards their social environment through these apps. Therefore, our first two hypotheses are:

H1.1: Acceptance values for data sharing with optional recipients are higher in T2 than in T1.

H1.2: Acceptance values for data collection of optional information attributes are higher in T2 than in T1.

5.2.3.2 Data recipients

Concerning potential data recipients, Apthorpe et al. (2018) use the immediate family members as baseline and argue that immediate family members usually have knowledge about the information that can be transmitted in their IoT scenarios. For example, immediate family members know from each other which travel vehicles are used or whether they do sports. Consistent with this argumentation, the authors observe the highest acceptance values for immediate family members. For the scenarios in our survey, we also assume that the highest acceptance values for data sharing are indicated for the recipient *household members*. On the contrary, we expect the lowest acceptance values to be indicated for the recipient *employer*. Persson and Hansson (2003) discuss several reasons why employers may have an interest in

invading the privacy of their employees. At the same time, employees can have numerous reasons why they would not want to share private information such as information collected in apps with their employers. In a study on the rating environment and platform design of employer review platforms, Cloos (forthcoming) argues that employees, due to their economic dependence on employers, have strong incentives to refrain from too permissive data sharing on the internet. For the scenarios in our survey, it also seems plausible that participants fear negative consequences if data is shared with employers. Although numerous negative consequences are also conceivable when data is shared with other recipients, these consequences are less dramatic than a possible job loss which might be the result of data sharing with employers. Therefore, our next hypotheses are:

H2.1: Acceptance values for *household members* are higher than acceptance values for the other optional recipients.

H2.2: Acceptance values for *employer* are lower than acceptance values for the other optional recipients.

5.2.3.3 Information attributes

Unlike for recipients, our scenarios do not include information attributes that are queried in all scenarios. When formulating the hypotheses on the acceptance towards data collection of optional information attributes, we exclusively concentrate on the *live location* and the *monthly net income*, since these information attributes are included in more than half of the scenarios. Live location data raise (serious) security and privacy concerns (Minch, 2015). In Apthorpe et al.'s (2018) study, the sharing of the live location by different IoT transmitters is evaluated as relatively unacceptable. The results of a qualitative study by Muslukhov et al. (2012) also show that smartphone users perceive location tracking as very sensitive. In line with this literature, we expect that the acceptance values regarding a transmission of the live location are very low in our study. With regard to data collection of the *monthly net income*, we also expect very low acceptance values. In Germany, it is relatively unusual to talk about one's own income. People tend to not want to talk about their own income and also feel that they should not talk about it (for a discussion of related surveys, see Sauerland and Höhs (2019) (in German)). Hence, we derive the following hypotheses:

H3.1: Acceptance values for *live location* and *monthly net income* are lower than acceptance values for the other optional information attributes.

H3.2: Acceptance values do not differ between *life location* and *monthly net income*.

5.2.3.4 Attitudes

The fourth set of hypotheses refers to respondents' attitudes on privacy, risk, and green consumption. To measure the privacy preferences of our participants, we use the scale by Ham (2017). The items of this scale ask for consent to collect data on participants' online behavior. Since the questions in our study refer to a very similar subject area, we expect higher values on the privacy scale to be associated with lower acceptance values. With regard to participants' risk attitudes, we expect that a higher willingness to take risks is associated with higher acceptance values. In a study on the privacy paradox (i.e., a potential privacy intentions behavior gap), Norberg et al. (2007) find that a higher risk aversion is associated with a lower willingness to provide personal data. Further research (Fogel and Nehmad, 2009) shows that individuals who use social networks are more willing to take risks than individuals who do not use social networks. To the best of our knowledge, there are no studies that investigate the relationship between green consumption values and privacy preferences. Therefore, we deliberately choose to not formulate any hypothesis on green consumption values and consider our study to be explorative in this respect. Our two hypotheses on participants' attitudes therefore are:

H4.1: Higher values on the privacy concern scale are associated with lower acceptance values.

H4.2: A higher risk propensity is associated with higher acceptance values.

5.2.3.5 Demographics

In terms of demographics, we investigate hypotheses on age, gender, and education. Goldfarb and Tucker (2012) find that older people are less willing to provide information on their own income in an online survey compared to younger people. Andone et al. (2016) investigate the smartphone usage behavior of different age groups based on tracking a sample of more than 30,000 participants for at least 28 days. Their results show that younger people use their smartphones more time intensively and with a larger number of specialized apps than older people. Based on this literature, we assume that younger participants are more open towards the app scenarios described in our survey and therefore indicate higher acceptance values. In an experiment on the willingness to disclose different types of personal information in exchange for money, Benndorf and Normann (2018) find that female participants mostly request significantly more money than male participants. Research on privacy preferences in social networks shows that while women and men share similar amounts of information privately with friends, men are significantly more willing to share information publicly

(Quercia et al., 2012). For our study, we therefore expect men to indicate higher acceptance values for data sharing in the scenarios described. In a national phone survey, Turow et al. (2005) examine the knowledge of 1,500 US Americans regarding data collection and data usage practices of commercial websites. The results show that the number of correctly answered questions was higher for participants with higher formal education. In this study, we therefore assume that participants with a higher formal education have more knowledge about data protection on the internet and, hence, are more skeptical about the scenarios described. Our hypotheses on participants' demographic attitudes are:

H5.1: A higher age is associated with lower acceptance values.

H5.2: Male participants have higher acceptance values.

H5.3: A higher formal education is associated with lower acceptance values.

5.2.4 Power analysis

In order to get an impression of the effect sizes at which we can detect significant treatment differences, we conducted a power analysis. For the power analysis, we estimated a mean value for the information attribute *live location* as an example, since a similar attribute ("its owners location") is also included in the study by Apthorpe et al. (2018). The authors use an acceptance scale that ranges from -1.5 to 1.5. In their study, the acceptance values for "its owners location" range from -0.67 to -0.28 with a mean value of -0.43 for various IoT senders. Transferred to our acceptance scale, which ranges from -1 to 1, this corresponds to an acceptance value of -0.29. However, since in all of our scenarios more attributes are transmitted than in the scenarios of Apthorpe et al. (2018), and since the senders in our scenarios are not IoT devices but actors from the private or public sector, we expect a slightly lower acceptance value, which we assume to be -0.4. Since no standard deviations are reported in Apthorpe et al. (2018), we assume a standard deviation of 0.5. In line with our hypothesis, we expect higher acceptance values in T2. Based on the power analysis, we estimate the minimum distance between the mean acceptance values that is required to obtain a significant result by using a two sample (one-sided) means test. Since we perform 30 (21) pairwise tests on hypothesis H1.1 (H1.2), there is a high probability for the occurrence of Type I errors. We therefore choose a low significance level of alpha = 0.005. For a significance level of alpha = 0.005, a power of 0.8, and a standard deviation of 0.5, the distance would have to be 0.107 (or 0.214 standard deviations) when considering a single scenario with an average participant number of N = 506

per treatment. We consider this calculated necessary effect size between the means of the two treatments to be large enough to indicate meaningful treatment effects.

5.2.5 Procedures

Our online survey was programmed with the software LimeSurvey. Before the survey was dispatched to participants, a pre-test for comprehensibility and length was conducted with six university researchers. In addition, we used the pre-test for a qualitative check of our scenarios in terms of the validation criteria of internal consistency, plausibility, creativity, and relevance (Amer et al., 2013). As a result, we refined the wording in some scenarios and included additional recipients.

The recruiting of the participants was conducted by a panel provider in September 2020. The participant sample is representative for the German population between 18 and 69 years in terms of gender, age, federal state, and education (see Appendix B (section 5.6.2)). The respondents' payment (50 cents per participant) was also processed via the panel provider. The email announced a "survey on digital technologies" to avoid a link to privacy (or data protection) research.

A total of 1,357 people participated in the survey. 10.61% (N = 144) of the respondents did not complete the questionnaire. Among the persons who answered the questionnaire completely, 14.01% (N = 170) did not pass the quality check question[37]. Accordingly, the sample size reduced to 1,043. In a second step, we identified and eliminated speeders. The median time to complete the questionnaire was 9:30 minutes. Participants (N = 23) requiring less than 1/3 of the interview time (3:10 minutes) were dropped. In a last step, seven people were removed after a manual quality check.[38] The final sample included 1,013 participants with a female share of 51.73% (N = 524) and an average age of 45.81 years (sd = 14.42).

5.3 Results

This section presents the results of our survey. In section 5.3.1, the acceptance values for each scenario are presented and possible treatment effects are examined. We deliberately avoid

[37] The quality check question was integrated as one item in the GREEN consumption scale. The wording of the question was: "To make sure that you read the questionnaire carefully, please select the answer option 'strongly agree (7)'."

[38] These participants showed no variance in their responses regarding the GREEN consumption scale and the privacy concerns scale, although reverse items were included. For these items, the respective participants always chose the answer option (7) "strongly agree", so that the quality check was randomly passed.

comparing acceptance values for identical recipients and information attributes in different scenarios since the different scenarios describe various sending device and provider combinations as well as varying mandatory data specifications and transmission benefits. Thereafter, we examine to what extent the acceptance values within the scenarios differ between recipients (section 5.3.2) and information attributes (section 5.3.3). Based on a regression analysis, we further investigate how the response behavior of our participants is influenced by their attitudes as well as their demographic characteristics (section 5.3.4).

5.3.1 Acceptance values and treatment effects

Tables 5.2 and 5.3 show the average acceptance values for optional recipients and optional information attributes in each scenario and for both treatments. Overall, in both tables, not a single acceptance value is greater than zero and therefore all values are in the unacceptable range. In Table 5.2 (recipients), the acceptance values range from -0.79 to -0.18. Similarly, in Table 5.3 (information attributes), the acceptance values range from -0.80 to -0.18. As depicted in the histograms in Appendix C (section 5.6.3), the relatively low average acceptance values can be explained by the fact that for each question a large number of participants chose the answer with the lowest value ("completely unacceptable"). A total of 10.86% (N = 110) chose this answer for each individual question in the scenarios A to E. In Tables 5.2 and 5.3, we do not observe a significant treatment effect for any of the different recipients and information attributes (pairwise comparisons with two-sample Wilcoxon rank-sum tests; all p-values > 0.005). In addition, there are no indications that acceptance values in one treatment are systematically higher or lower than in the other treatment. We thus reject hypotheses H1.1 and H1.2.

Scenario	A		B		C		D		E			
Sending device and provider	Loyalty card with app from supermarket chains		Tracking bracelet with app from a health insurance company		Nutrition app from the Federal Ministry of Health		Mobility tracking app from a start-up company		Smart meter with app from an energy provider		Average acceptance values	
Recipient	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2
Household members	-.20 (.70)	-.21 (.70)	-.29 (.71)	-.34 (.69)	-.38 (.68)	-.42 (.68)	-.39 (.68)	-.42 (.68)	-.21 (.72)	-.20 (.71)	-.29 (.60)	-.32 (.58)
Employer	-.66 (.55)	-.68 (.56)	-.69 (.52)	-.71 (.53)	-.74 (.47)	-.76 (.47)	-.76 (.47)	-.77 (.46)	-.79 (0.42)	-.78 (.45)	-.73 (.43)	-.74 (.44)
Federal ministries	-.36 (.66)	-.42 (.62)	-.44 (.63)	-.47 (.63)	-	-	-.52 (.60)	-.53 (.61)	-.51 (.60)	-.51 (.60)	-.46 (.53)	-.48 (.54)
City or municipality	-	-	-	-	-	-	-.60 (.56)	-.57 (0.60)	-.51 (.59)	-.47 (.61)	-.55 (.52)	-.52 (.55)
Market research comp.	-.18 (.67)	-.19 (.66)	-.38 (.65)	-.39 (.65)	-.44 (.62)	-.43 (.65)	-	-	-	-	-.33 (.58)	-.34 (.58)
German food prod.	-.26 (.65)	-.26 (.66)	-	-	-.55 (.57)	-.49 (.63)	-	-	-	-	-.40 (.55)	-.37 (.58
American food prod.	-.65 (.52)	-.63 (.54)	-	-	-	-	-	-	-	-	-.65 (.52)	-.63 (.54)
Chinese food prod.	-.69 (.51)	-.70 (.51)	-	-	-	-	-	-	-	-	-0.69 (0.51)	-0.70 (0.51)
German sports equip. prod.	-	-	-.53 (.60)	-.54 (.60)	-	-	-	-	-	-	-.53 (.60)	-.54 (.60)
German elect. prod.	-	-	-	-	-	-	-	-	-.51 (.60)	-.48 (.61)	-.51 (.60)	-.48 (.61)
Health insurance company	-.48 (.63)	-.48 (.65)	-	-	-	-	-	-	-	-	-.48 (.63)	-.48 (.65)
Car insurance company	-	-	-	-	-	-	-.65 (.52)	-.67 (.53)	-	-	-.65 (.52)	-.67 (.53)
Public transport companies	-	-	-	-	-	-	-.51 (.61)	-.52 (.62)	-	-	-.51 (.61)	-.52 (.62)
German Society for Nutrition	-	-	-	-	-.41 (.64)	-.40 (.66)	-	-	-	-	-0.41 (0.64)	-0.40 (0.66)
Neighbors	-	-	-	-	-	-	-	-	-.78 (.42)	-.78 (.45)	-.78 (.42)	-.78 (.45)
Average acceptance values	-.44 (.47)	-.45 (.49)	-.47 (.52)	-.49 (.52)	-.50 (.51)	-.50 (.53)	-.57 (.48)	-.58 (.48)	-.55 (.44)	-.54 (.46)		
$n =$	505	508	505	508	505	508	505	508	505	508	505	508

Average acceptance values: Color division					
	-1 to -0.8	> -0.8 to -0.6	> -0.6 to -0.4	> -0.4 to -0.2	> -0.2 to 0
	> 0 to 0.2	> 0.2 to 0.4	> 0.4 to 0.6	> 0.6 to 0.8	> 0.8 to 1

Table 5.2: Mean acceptance values for different data recipients in scenarios A to E.

Notes. The row 'average acceptance values' refers to the average acceptance value for all recipients included in the respective scenario. The column 'average acceptance values' refers to the average acceptance value for all scenarios where the respective recipient is included. Acceptance values for specific recipients are never significantly different between T1 and T2 when assuming a significance level of $p < 0.005$ (two-sample Wilcoxon rank-sum test). Even at a higher significance level of $p < 0.05$, none of the differences is significant.

Scenario	A		B		C		D		E			
Sending device and provider	Loyalty card with app from supermarket chains		Tracking bracelet with app from a health insurance company		Nutrition app from the Federal Ministry of Health		Mobility tracking app from a start-up company		Smart meter with app from an energy provider		Average acceptance values	
Attribute	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2
Live location	-.60 (.56)	-.60 (.56)	-.61 (.57)	-.62 (.56)	-.68 (.52)	-.70 (.51)	-	-	-.67 (.54)	-.71 (.51)	-.64 (.49)	-.66 (.47)
Monthly net income	-	-	-	-	-.73 (.48)	-.75 (.49)	-.76 (.46)	-.77 (.45)	-.75 (.47)	-.80 (.43)	-.75 (.44)	-.78 (.42)
Nutrition style	-.29 (.66)	-.21 (.69)	-.37 (.68)	-.33 (.69)	-	-	-	-	-	-	-.33 (.62)	-.27 (.63)
Food intol. and allergies	-.25 (.67)	-.18 (.72)	-	-	-.34 (.69)	-.32 (.72)	-	-	-	-	-.30 (.62)	-.25 (.65)
Body weight and height	-.53 (.59)	-.48 (.62)	-.37 (.67)	-.36 (.69)	-	-	-	-	-	-	-.45 (.57)	-.42 (.60)
Member-ships in a sports club or gym	-	-	-.44 (.64)	-.43 (.66)	-.52 (.60)	-.53 (.62)	-	-	-	-	-.48 (.59)	-.48 (.60)
Profession	-	-	-	-	-	-	-.55 (.59)	-.55 (.61)	-.64 (.54)	-.65 (.56)	-.60 (.52)	-.60 (.55)
Number of steps taken	-.45 (.62)	-.43 (.65)	-	-	-	-	-	-	-	-	-.45 (.62)	-.43 (.65)
Date of birth	-	-	-	-	-	-	-.40 (.65)	-.41 (.66)	-	-	-.40 (.65)	-.41 (.66)
Driving behavior	-	-	-	-	-	-	-.49 (.61)	-.50 (.63)	-	-	-.49 (.61)	-.50 (.63)
Time and duration of use of indi. power sources	-	-	-	-	-	-	-	-	-.26 (.68)	-.25 (.69)	-.26 (.68)	-.25 (.69)
Average acceptance values	-.42 (.51)	-.38 (.53)	-.45 (.56)	-.43 (.56)	-.57 (.49)	-.58 (.49)	-.55 (.50)	-.56 (.50)	-.58 (.45)	-.60 (.43		
$n =$	505	508	505	508	505	508	505	508	505	508	505	508

Average acceptance values: Color division					
	-1.0 to -.8	> -.8 to .6	> -.6 to -.4	> -.4 to -.2	> -.2 to 0.0
	> 0.0 to .2	> .2 to .4	> .4 to .6	> .6 to .8	> .8 to 1.0

Table 5.3: Mean acceptance values for different information attributes in scenarios A to E.

Notes. The row 'average acceptance values' refers to the average acceptance value for all information attributes included in the respective scenario. The column 'average acceptance values' refers to the average acceptance value for all scenarios where the respective information attribute is included. Acceptance values for specific information

attributes are never significantly different between T1 and T2 when assuming a significance level of $p < 0.005$ (two-sample Wilcoxon rank-sum test). Even at a higher significance level of $p < 0.05$, none of the differences is significant.

5.3.2 Acceptance towards data sharing with optional data recipients

In this section, differences in the acceptance towards data sharing with optional recipients are investigated for each scenario. For space reasons, we will not discuss each individual result in detail. Instead, we focus on results that are related to the hypotheses H2.1 and H2.2. We use non-parametric Wilcoxon signed-rank tests (matched samples) to examine whether acceptance values differ between optional data recipients.[39] Since we do not find any treatment effects and our hypotheses on optional recipients do not refer to individual treatments, we use pooled data for the pairwise tests. The results for all pairwise tests can be found in Appendix D (section 5.6.4), Tables 5.11 to 5.15.

As Table 5.2 shows, comparatively high acceptance values can be observed for *household members* (ranging from -0.42 in scenarios C and D to -0.20 in scenario A) while *employers* belong to the recipients with the lowest acceptance values across all scenarios (ranging from -0.79 in scenario E to -0.66 in scenario A). In all scenarios the acceptance values for data sharing with *household members* are almost always significantly higher than for other recipients (Wilcoxon signed rank tests, p-values < 0.01). In scenario A (Appendix D, Table 5.11), however, there is no significant difference between *household members* and *market research companies* (p-value = 0.633). In scenario C (Appendix D, Table 5.13), the difference is also not significant for *market research companies* (p-value = 0.014) and for the *German Society for Nutrition* (p-value = 0.877). Concerning data sharing with *employers*, the acceptance values are almost always significantly lower than for other recipients (p-values < 0.01), except for *American food producers* (p-value = 0.306) and *Chinese food producers* (p-value = 0.067) in scenario A (Appendix D, Table 5.11), and for *neighbors* (p-value = 0.431) in scenario E (Appendix D, Table 5.15). We thus find predominant support for hypotheses H2.1 and H2.2.

5.3.3 Acceptance towards data collection of optional information attributes

This section focuses on differences in the acceptance values for data collection of optional information attributes. In order to test hypotheses H3.1 and H3.2, we compare acceptance values from all scenarios which include the optional information attributes *life location* and

[39] In sections 5.3.2 and 5.3.3, we use non-parametric Wilcoxon signed-rank tests instead of parametric t-tests because the differences between the individual acceptance values are not normally distributed.

monthly net income. Tables 5.16 to 5.20 in Appendix D (section 5.6.4) show the results for all pairwise tests.

In accordance with hypotheses H3.1, we observe the lowest acceptance values for the information attributes *live location* (ranging from -0.71 in scenario E to -0.60 in scenario A, see Table 5.3) and *monthly net income* (ranging from -0.80 in scenario E to -0.73 in scenario C). The acceptance values for *live location* are significantly lower than the acceptance values for all other information attributes in scenario A and B (Appendix D, Tables 5.16 and 5.17, all p-values < 0.001). Likewise, the acceptance values for *monthly net income* are significantly lower than the acceptance values for all other information attributes in scenarios C, D, and E (Appendix D, Tables 5.18, 5.19 and 5.20, all p-values < 0.001). Therefore, we accept hypothesis H3.1. In scenarios C and E, the acceptance values for *monthly net income* are significantly lower than for *live location* (both p-values < 0.001). We thus, reject hypothesis H3.2.

5.3.4 The influence of personal attitudes and demographics

In this section, we use random-effects generalized least squares (GLS) regression models to test our hypotheses on attitudes and demographics. In Table 5.4, the dependent variable is either the acceptance value indicated for each data recipient (models 1-4) or for each information attribute (models 5-8). Since each participant indicated a total of 30 acceptance values for data recipients and a total of 21 acceptance values for information attributes, the number of observations in Table 5.4 is $1{,}013 * 30 = 30{,}390$ in models 1-4 and $1{,}013 * 21 = 21{,}273$ in models 5-8.

The upper three independent variables in Table 5.4 are related to participants' personal attitudes. *Privacy* and *GREEN* are scores calculated from the average answer values to the privacy concern scale of Ham (2017) and the GREEN scale of Haws et al. (2014). For the privacy scale, we obtain a Cronbach's alpha of 0.69 and for the GREEN scale a Cronbach's alpha of 0.71. Mean values and standard deviations for each item of these scales can be found in Appendix E (Tables 5.21 and 5.23). The variable *Risk* contains the indicated risk propensity of the participants based on Dohmen et al. (2011) where a higher number indicates a higher willingness to take risks. The mean value and standard deviation for this question can be found in Appendix E (Table 5.22). *Education (high)* and *Education (low)* are dummy variables that

take a value of 1 if the participant has a high school degree, or a degree from a basic secondary school or lower.[40]

Models 1 and 5 analyze the influence of participants' online privacy concerns, risk propensity, and green consumption values on the acceptance towards data sharing (model 1) and data collection (model 5). In models 2 and 6, we focus on the influence of demographic variables. In models 3 and 7, we consider both personal attitudes and demographic characteristics. Since previous research by Dohmen et al. (2017, 2011) shows that risk attitudes are higher for younger people and for men, we include the interaction terms *Risk*Age* and *Risk*Male* in models 4 and 8. The results of the studies by Fast and Schnurr (2020) and Fogel and Nehmad (2009) further show that women, on average, have higher privacy concerns than men. We therefore include the interaction term *Privacy*Male* in models 4 and 8. All regression models contain control dummies for optional recipients (models 1-4) or for optional information attributes (models 5-8).

[40] The variable Education (medium) is omitted.

	Dep. var. acceptance values for optional							
	data recipients				**information attributes**			
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Privacy	-0.17***		-0.16***	-0.12***	-0.17***		-0.17***	-0.13***
	(0.01)		(0.01)	(0.02)	(0.01)		(0.01)	(0.02)
Risk	0.02***		0.01*	0.04*	0.02***		0.01*	0.03
	(0.01)		(0.01)	(0.02)	(0.01)		(0.01)	(0.02)
GREEN	0.03*		0.04**	0.03**	0.03*		0.04**	0.04**
	(0.01)		(0.01)	(0.01)	(0.01)		(0.01)	(0.01)
Age		-0.01***	-0.01***	-0.00		-0.01***	-0.01***	-0.00
		(0.00)	(0.00)	(0.00)		(0.00)	(0.00)	(0.00)
Gender (Male=1)		0.05	0.02	0.29*		0.06*	0.03	0.34*
		(0.03)	(0.03)	(0.13)		(0.03)	(0.03)	(0.14)
Education (high)		-0.10**	-0.09**	-0.09**		-0.06	-0.05	-0.05
		(0.03)	(0.03)	(0.03)		(0.03)	(0.03)	(0.03)
Education (low)		0.00	-0.01	-0.01		-0.01	-0.02	-0.03
		(0.03)	(0.03)	(0.03)		(0.03)	(0.03)	(0.03)
Risk*Age				-0.00**				-0.00*
				(0.00)				(0.00)
Risk*Male				0.03**				0.03**
				(0.01)				(0.01)
Privacy*Male				-0.08**				-0.08***
				(0.02)				(0.02)
Treatment (T2=1)	0.00	-0.01	0.00	0.01	0.02	0.01	0.02	0.02
	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.02)
Constant	0.13	-0.13*	0.46***	0.06	0.00	-0.32***	0.31**	-0.06
	(0.09)	(0.06)	(0.10)	(0.14)	(0.09)	(0.06)	(0.10)	(0.15)
Controls for optional recipients / optional information attributes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	30390	30390	30390	30390	21273	21273	21273	21273
N (groups)	1013	1013	1013	1013	1013	1013	1013	1013
Wald Chi^2	5234.68	5100.38	5311.24	5349.44	3813.26	3652.12	3872.95	3905.44

Table 5.4: Random-effects GLS regression of acceptance towards data sharing with optional recipients (models 1 4) or of acceptance towards data collection of optional information attributes (models 5-8) on participants' attitudes, demographic characteristics, and treatment.

Notes. Standard errors in parentheses: * $p < 0.01$, ** $p < 0.005$, *** $p < 0.001$.

Our results show significant negative effects for the variable *Privacy*. Stronger preferences for privacy have a large negative effect on the acceptance towards data sharing and on the acceptance towards data collection. This result is not surprising since the dependent variables in Table 5.4 and the privacy variable measure, in a broad sense, similar outcomes. Therefore, we accept hypothesis H4.1. In line with hypothesis H4.2, *Risk* has a significant positive effect on the dependent variable except for model 8. This suggests that risk takers are, on average, less reluctant to disclose their personal information. We did not formulate a hypothesis for a possible effect of the *GREEN* scale due to a lack of appropriate literature. In all models that include the GREEN scale, a higher value on the *GREEN* scale has a significant positive effect on the dependent variable. This implies that participants with a higher value on the *GREEN* scale are, on average, more open to disclose their personal information

For the demographic variables, we observe a significant negative effect of the variable *Age* in models 2, 3, 6, and 7. In models 4 and 8, *Age* is no longer significant, but the interaction term of the variables *Age* and *Risk* is significant and negative. The interaction term indicates that the significant negative effect of Age is mainly driven by the higher risk aversion of older participants. Thus, we find predominant support for hypothesis H5.1. The variable *Male* has a significant positive effect in models 4, 6, and 8. We therefore accept hypothesis H5.2. The interaction terms *Risk*Male* and *Privacy*Male* suggest that the effects of a stronger risk propensity and of stronger privacy concerns on the dependent variables significantly differ between women and men. For men, a higher value on the risk propensity (privacy concern) scale is associated with significantly higher (lower) acceptance values. In models 2, 3, and 4, *Education (high)* has a significant negative effect on the dependent variable. We do not observe significant effects of this variable for models 5 to 8 where the acceptance towards the collection of optional information attributes is the dependent variable. However, since the sign of *Education (high)* is also negative in models 5 to 8, we accept hypothesis H5.3.

5.4 Discussion and limitations

The results from Tables 5.2 and 5.3 show that our two different treatments do not lead to significant differences in the average acceptance values for the same recipients or information attributes. In addition, no clear trend can be identified.

The treatments in our scenarios differed only in one sentence, which emphasized different transmission benefits at the end of each scenario. Therefore, a possible explanation for the non-existent or only minor treatment effects is that the emphasized transmission benefits between the two treatments did not differ sufficiently. It is also possible that some of the participants interpreted the environmental benefits highlighted in T2 as not trustworthy. The results of a study conducted with Portuguese students by Do Paço and Reis (2012), show that students with particularly strong environmental concerns tended to be particularly skeptical about environmentally friendly advertising messages from companies. The mean across all items of the GREEN scale suggests that participants in our survey have comparatively strong environmental concerns. In Haws et al. (2014), the mean of the GREEN scale is 3.95, whereas for our participants it is significantly higher with 4.71. It is likely that participants in T2 in our survey were skeptical about the highlighted environmental benefits of the app and that these benefits therefore had no effect on the indicated acceptance values. In a study on greenwashing (i.e., deceptive advertising about the environmental characteristics of goods),

Schmuck et al. (2018) find that the negative effect of perceived greenwashing statements can be outweighed by nature images presented together with the greenwashing statements. It is thus also possible that a more detailed description or a visual presentation of the respective transmission benefits in our survey would have resulted in more pronounced treatment effects. A further explanation for the lack of treatment effects is that the acceptance decisions queried in the scenarios were simply less influenced by the transmission benefits but more by the perceived threats of data sharing.

The average acceptance values of the individual scenarios in our survey are not directly comparable since the scenarios differ in several parameters. Nevertheless, in each scenario, special care was taken to include both relatively uncritical as well as sensitive recipients and information attributes. The results show that the acceptance values are highest in the app scenarios that are probably relatively familiar to the participants. As explained in section 5.2.1.1, apps similar to those in scenario A (loyalty card) and B (tracking bracelet) also have a significantly higher market penetration than apps similar to those in scenario D (mobility tracking) or E (smart meter). Smart meters will be mandatory in all German households by the year 2032. Therefore, energy providers and relevant public authorities can use the low acceptance values in scenario E (smart meter) as an indication that broad-based information campaigns may be necessary to increase acceptance of this technology.

In general, we observe that the comparatively highest acceptance values were indicated for information attributes that show a close thematic fit with the respective scenario. For example, comparatively high acceptance values were indicated for *nutrition style* and *food intolerances and allergies* in scenario A (loyalty card). It is quite plausible that data on these information attributes can be used to make the personalized product recommendations described in scenario A as accurate as possible. This is not the case for the information attributes *body weight and height* and *number of steps taken*, which have significantly lower acceptance values. Another example can be found in scenario E (smart meter). Here, the acceptance value for the information attribute *time and duration of use of the individual power sources* is significantly higher than for *profession*.

In our regression analysis (Table 5.4), we found a significant positive effect of the GREEN scale on the acceptance values. One explanation for this effect could be that participants with higher values on the GREEN scale assume that they behave in accordance with existing social norms on environmental aspects. Those participants may be less concerned that the disclosure

of personal information may have negative consequences for them and therefore chose higher acceptance values in our survey.

One limitation of this study lies in the selection of the participants. Although the respondents were selected according to quotas for gender, age, federal state, and education, it can be assumed that the participants of our survey do not fully represent the German population. Since our study was conducted via a professional panel provider and with comparatively low monetary incentives, it is likely that our participants have an above-average internet affinity and intrinsic motivation. This assumption is further supported by the fact that participants accepted the invitation for an online survey on digital technologies.

A second limitation is that our study did not evaluate actual data sharing behavior. It is quite likely that in reality, participants of our survey would share personal data without much concern, even though they indicated low levels of acceptance in our survey. In economic experiments, in which participants decide on actual payoff relevant actions, it is often the case that participants show a comparatively open data transfer behavior, although they previously stated strong privacy preferences (Kokolakis, 2017; Norberg et al., 2007). Therefore, our results cannot be used to draw direct conclusions about participants' actual data sharing behavior. However, there is no reason to assume that the differences between different recipients and information attributes and the effects of attitudes and demographic characteristics revealed in our results are not reflected in real world situations.

5.5 Conclusion and outlook

The aim of our study was to examine whether and how the acceptance regarding data sharing in smartphone apps from five different industries differs for several data recipients and information attributes. In two treatments, we further investigated whether acceptance values are higher when environmental (public) instead of monetary (private) data transmission benefits are highlighted. Results show no treatment effects for data sharing with different recipients and for collection of various information attributes.

Our results show statistically significant differences in acceptance values between almost all recipients and between almost all information attributes. Comparatively high acceptance values were identified for the recipients and information attributes that thematically corresponded with the respective scenario. In line with our hypotheses, comparatively high acceptance values were stated for the recipient *household members* while the lowest acceptance values were stated for

employers. For the information attributes, our results revealed the lowest acceptance values for *live location* and *monthly net income*. The results from a regression analysis showed that the participants' age, a higher education level, and strong privacy concerns had a significant negative effect on acceptance values. In contrast, we found that participants with stronger GREEN consumption values, a higher willingness to take risks, and male participants had, on average, higher acceptance values.

For developers and providers of technologies that may raise privacy concerns among potential users, our study provides illustrative examples on how to investigate acceptance toward the technology in question. Future research could examine whether differences in the acceptance evaluation of data sharing in smartphone apps (or stationary digital applications) become apparent when there is a more intensive and/or visual emphasis of monetary (private) and environmental (public) data sharing benefits. Scholars could further investigate to what extent the general acceptance of new technologies, which could be measured, e.g., with the Technological Readiness Adoption Index (Ramírez-Correa and Rondán-Cataluña, 2020), is affected by the privacy preferences of potential technology adopters. Within institutional and health economics, scenario-based approaches similar to those in this study could be used to ex ante evaluate public acceptance toward planned policies. Future studies could further use scenario-based surveys to examine the effect of design changes on, e.g., employer review platforms (Cloos, forthcoming) or online marketplaces.

Acknowledgements

We thank Matthias Greiff, Rainer Kühl, Christoph Bühren, Merindah Loessl, Jörg Müller and Laurin Spahn as well as the participants of the Clausthaler Ökonomisches Oberseminar for helpful feedback and comments. Jenny Andris provided excellent research assistance.

5.6 Appendix

5.6.1 A – Scenarios

In all scenarios answer options were: (1) Completely unacceptable; (2) Somewhat unacceptable; (3) Rather unacceptable; (4) Rather acceptable; (5) Somewhat acceptable; (6) Completely acceptable.

Scenario A - Loyalty card with app from supermarket chains	
Various supermarket chains collect information about the food you buy with a free, shared loyalty card in combination with a cost-free app. **Name**, **address**, **date of birth, and gender** must be entered in the app. Based on this information, the app will provide you with product recommendations tailored to your shopping behavior.	
T1	**T2**
The app also offers discount coupons, which **help you to save money while shopping**.	The supermarket chains emphasize that no advertising leaflets are printed for their app users, thus **avoiding waste and protecting the environment.**
Recipients	
How acceptable is it to you that the information collected with the loyalty card, is passed on to the following recipients in addition to the supermarket chains? The data transmission is mentioned in the app's general terms and conditions.	
1. Market research companies	
2. German food producers	
3. American food producers	
4. Chinese food producers	
5. Household members (close family members or roommates)	
6. Federal ministries (e.g., for health, economic affairs, environment, transport)	
7. Health insurance company	
8. Employer	
Information attributes	
How acceptable is it to you that the supermarket chains request or collect and store the following additional information from the app user(s) via the app if this is mentioned in the app's general terms and conditions?	
1. Live location	
2. Nutrition style (e.g., vegetarian/vegan/...)	
3. Number of steps taken	
4. Food intolerances and allergies	
5. Bodyweight and height	

Table 5.5: Description of scenario A.

Scenario B - Tracking bracelet with app from a health insurance company	
Your health insurance company provides you with a cost-free fitness bracelet in combination with a cost-free app to collect information about your activity status. Your daily steps and your heart rate form your activity status and determined via the fitness bracelet and automatically stored in the app. **Name, date of birth, and gender** must be entered in the app. The collected information is passed on to your health insurance company. Based on this information, your health insurance company will determine a weekly number of steps to be reached.	
T1	**T2**
For each week in which you reach the determined number of steps, **you will receive a bonus of € 3.00.**	For each week in which you reach the determined number of steps, **your health insurance company will assume a tree sponsorship of € 3.00 for worldwide reforestation projects.**
Recipients	
How acceptable is it to you that the information collected with the app is passed on to the following recipients, in addition to the health insurance company? The data transmission is mentioned in the app's general terms and conditions.	
1. Market research companies	
2. Household members (close family members or roommates)	
3. Employer	
4. Federal ministries (e.g., for health, economic affairs, environment, transport)	
5. German sports equipment producers	
Information attributes	
How acceptable is it to you that the health insurance company requests and respectively collects and stores the following additional information from the app user(s) via the app if this is mentioned in the app's general terms and conditions?	
1. Live location	
2. Nutrition style (e.g., vegetarian/vegan/...)	
3. Membership in a sports club or gym	
4. Body weight and height	

Table 5.6: Description of scenario B.

Scenario C – Nutrition app from the Federal Ministry of Health	
The Federal Ministry of Health offers a cost-free app to give you personalized nutritional recommendations. **Name, date of birth, gender, body weight, height, and your nutritional style** must be entered in the app. In addition, you have to provide information about your typical weekly purchases to the app by photographing the corresponding receipts at regular intervals. This information is passed on to the Federal Ministry of Health.	
T1	**T2**
Based on this information, the app will provide you with personalized nutritional recommendations **aimed at maximizing health-promoting nutrition**.	Based on this information, the app provides you with personalized nutritional recommendations **aimed at maximizing environmentally friendly nutrition**.
Recipients	
How acceptable is it to you that the information collected with the app is passed on to the following recipients, in addition to the Federal Ministry of Health? The data transmission is mentioned in the app's general terms and conditions of the app.	
1. Market research companies	
2. German food producers	
3. Household members (close family members or roommates)	
4. Employer	
5. German Society for Nutrition (independent scientific society)	
Information attributes	
How acceptable is it to you that the Federal Ministry of Health requests, or collects and stores the following additional information from the app user(s) via the app, if this is mentioned in the app's general terms and conditions?	
1. Live location	
2. Food intolerances and allergies	
3. Membership in a sports club or gym	
4. Monthly net income	

Table 5.7: Description of scenario C.

Scenario D - Mobility tracking app from a start-up company	
A German technology start-up company collects information about your mobility behavior via a cost-free tracking app. **Name, gender, live location, and type of vehicle** (assume you own a car) must be entered obligatorily for the app. The tracking app registers whether you travel by car, public transport, bicycle, or by foot. In addition, the app has information on the current location-based petrol, diesel and electricity prices, on the prices of public local and long-distance transport, and on the current traffic loads on roads and public transport.	
T1	**T2**
Based on this information, the app provides you with personalized recommendations **aimed at maximizing cost- and time-saving mobility behavior**.	Based on this information, the app provides you with personalized recommendations **aimed at maximizing environmentally friendly mobility behavior**.
Recipients	
How acceptable is it to you that the information collected with the app (except for live location data) is shared with the following recipients, in addition to the start-up company? The data transmission is mentioned in the app's general terms and conditions.	
1. Local and long-distance public transport companies	
2. City or municipality (residence)	
3. Household members (close family members or roommates)	
4. Employer	
5. Federal ministries (e.g., for health, economic affairs, environment, transport)	
6. Car insurance company	
Information attributes	
How acceptable is it to you that the start-up company requests, or collects and stores the following additional information from the app user via the tracking app, if this is mentioned in the app's general terms and conditions?	
1. Date of birth	
2. Driving behavior (when using the car as driver)	
3. Profession	
4. Monthly net income	

Table 5.8: Description of scenario D.

Scenario E – Smart meter with app from an energy provider	
In your apartment (or flat-sharing community, or house) a smart meter with connected measuring systems is installed. The smart meter is an intelligent digital electricity meter that records and stores your power consumption at any time and can send the stored data. The smart meter receives data from the connected measuring systems, which record the electricity consumption of individual power sources (e.g., tv, refrigerator, room lighting) in your apartment. Through an app of your energy provider, which receives data from your smart meter, information about your current and past electricity consumption is provided to you, clearly arranged by the power source. **Name**, **address, and date of birth** must be entered in the app.	
T1	**T2**
Based on this information, the app provides you with personalized recommendations **aimed at minimizing your electricity costs**.	Based on this information, the app provides you with personalized recommendations **aimed at maximizing environmentally friendly power usage**.
Recipients	
How acceptable is it to you that the information collected with the app is passed on to the following recipients, in addition to the energy provider? The data transmission is mentioned in the app's general terms and conditions.	
1. Local and long-distance public transport companies	
2. City or municipality (residence)	
3. Household members (close family members or roommates)	
4. Employer	
5. Neighbors	
6. Federal ministries (e.g., for health, economic affairs, environment, transport)	
Information attributes	
How acceptable is it to you that the energy provider requests and respectively collects and stores the following additional information of the app user(s) via the app if this is mentioned in the app's general terms and conditions?	
1. Live location	
2. Profession	
3. Monthly net income	
4. Time and duration of use of the individual power sources	

Table 5.9: Description of scenario E.

5.6.2 B - Demographics

Quote	Specification	N (%)
Gender	Male	489 (48.27)
	Female	524 (51.73)
Age	18-29	185 (18.26)
	30-39	178 (17.57)
	40-49	177 (17.47)
	50-59	262 (25.86)
	60-69	211 (20.83)
Education	Basic secondary schooling or lower	337 (33.27)
	Intermediate school certificate or equivalent	314 (31.00)
	High school graduation	362 (35.73)
Federal State	Baden Wurttemberg	132 (13.03)
	Bavaria	167 (16.49)
	Berlin	43 (4.24)
	Brandenburg	29 (2.86)
	Bremen	7 (0.69)
	Hamburg	23 (2.27)
	Hesse	77 (7.60)
	Lower Saxony	103 (10.17)
	Mecklenburg Western Pomerania	19 (1.88)
	North Rhine-Westphalia	215 (21.22)
	Rhineland-Palatinate	54 (5.33)
	Saarland	12 (1.18)
	Saxony	47 (4.64)
	Saxony-Anhalt	27 (2.67)
	Schleswig Holstein	35 (3.46)
	Thuringia	23 (2.27)

Table 5.10: Demographics of survey participants.

5.6.3 C – Histograms acceptances values

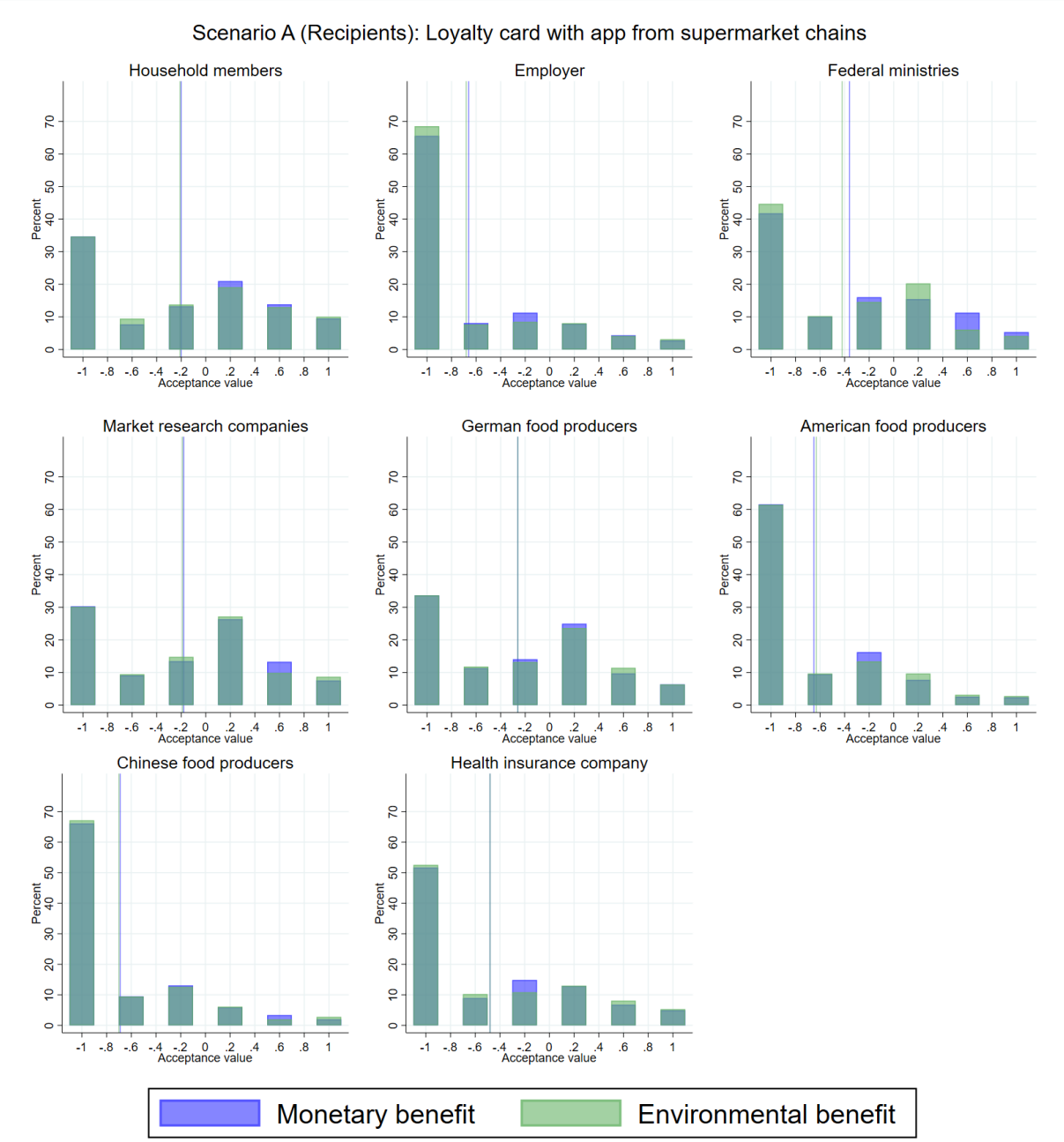


Figure 5.2: Histograms showing the percentage distribution of acceptance values for optional recipients in scenario A by treatment.

Notes. Vertical lines indicate the mean values in T1 (blue) and T2 (green).

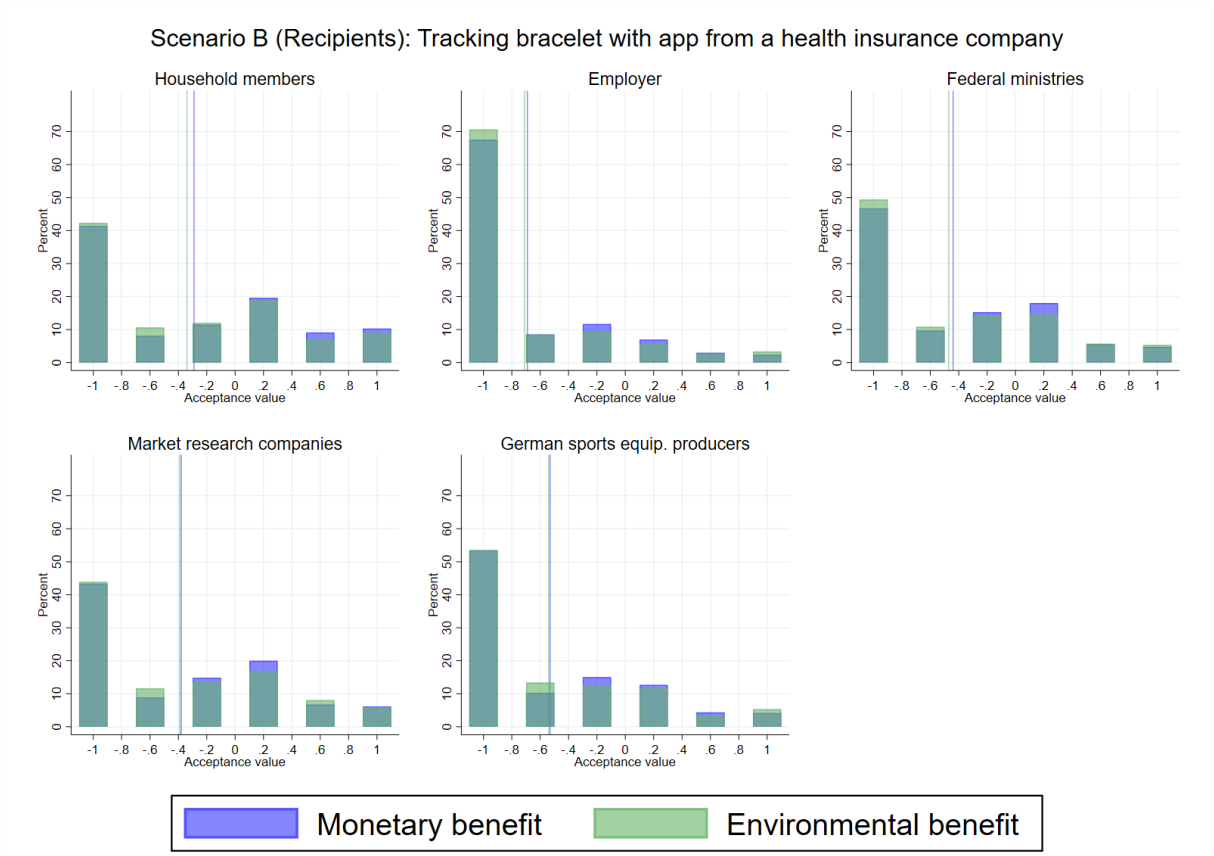


Figure 5.3: Histograms showing the percentage distribution of acceptance values for optional recipients in scenario B by treatment.

Notes. Vertical lines indicate the mean values in T1 (blue) and T2 (green).

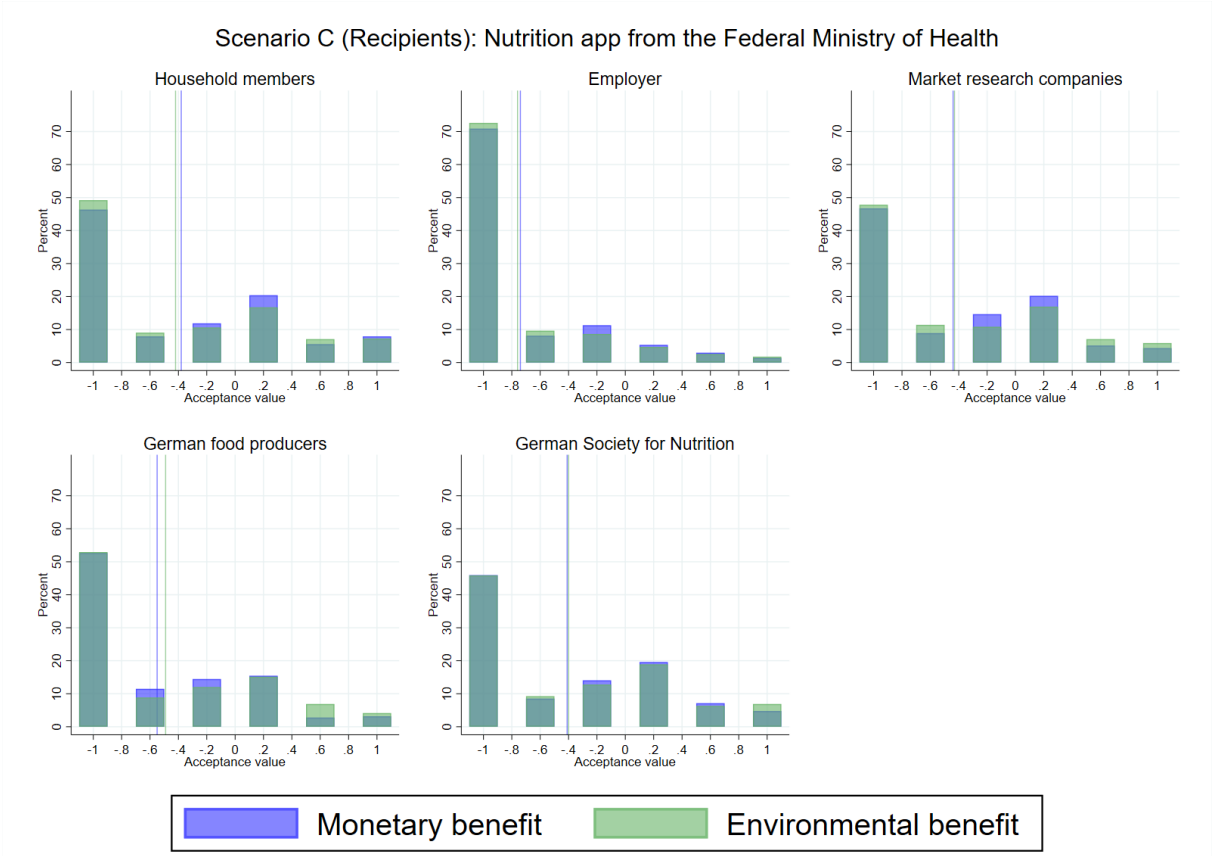


Figure 5.4: Histograms showing the percentage distribution of acceptance values for optional recipients in scenario C by treatment.

Notes. Vertical lines indicate the mean values in T1 (blue) and T2 (green).

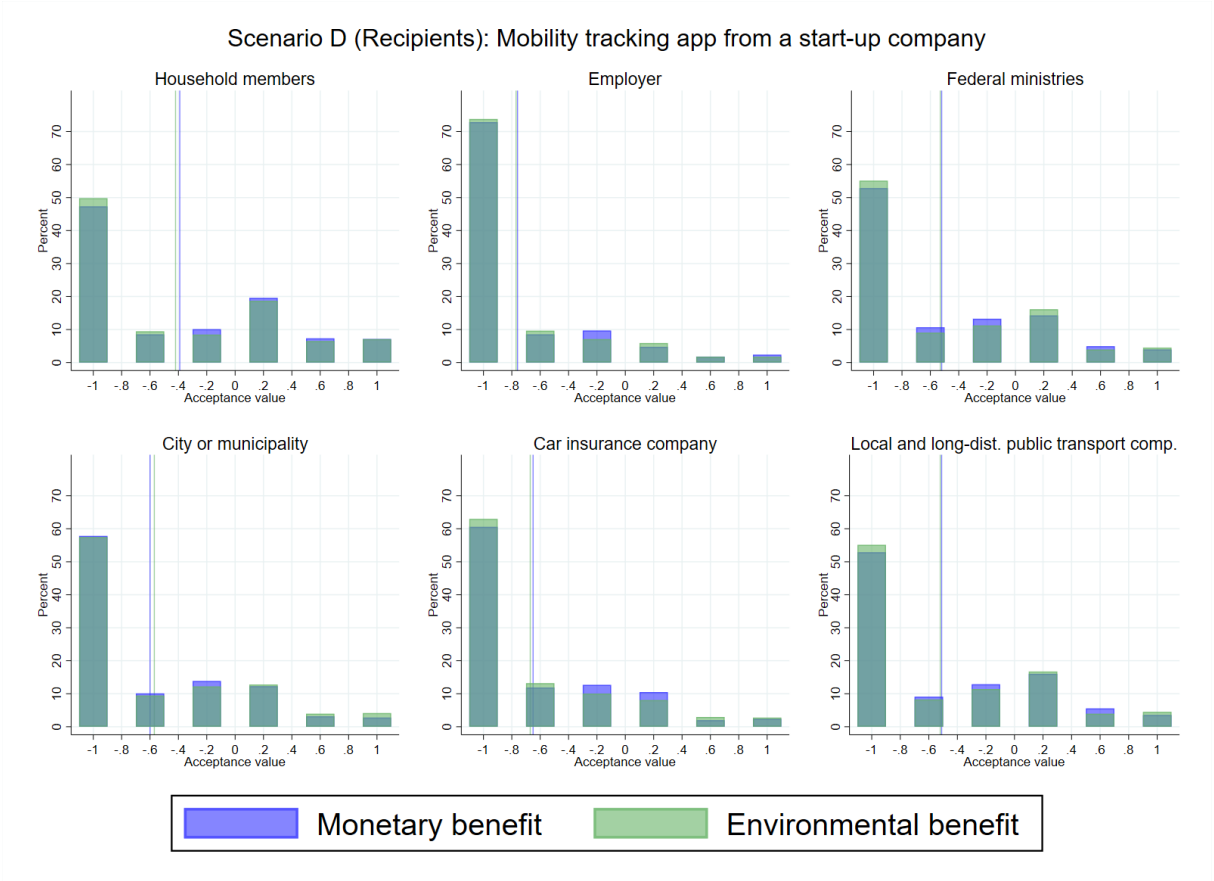


Figure 5.5: Histograms showing the percentage distribution of acceptance values for optional recipients in scenario D by treatment.

Notes. Vertical lines indicate the mean values in T1 (blue) and T2 (green).

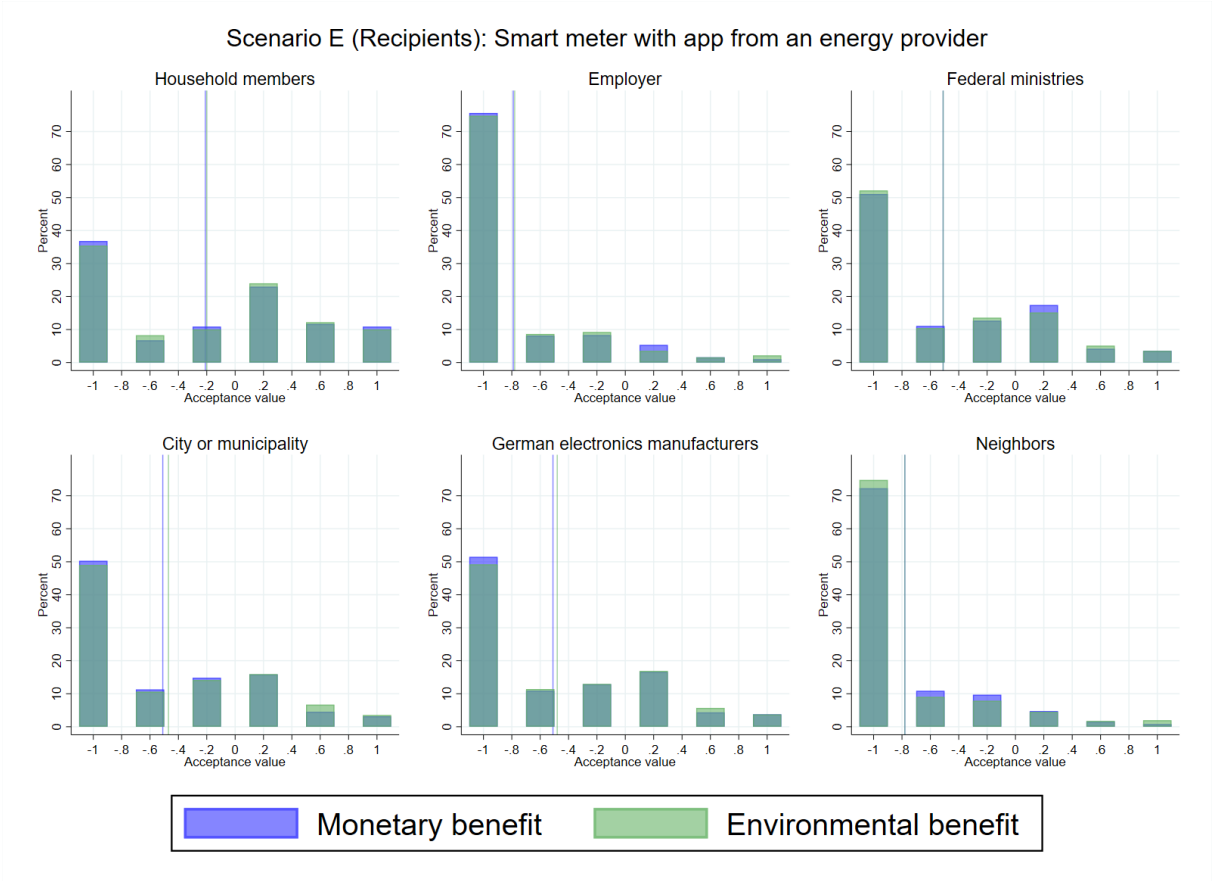


Figure 5.6: Histograms showing the percentage distribution of acceptance values for optional recipients in Scenario E by treatment.

Notes. Vertical lines indicate the mean values in T1 (blue) and T2 (green).

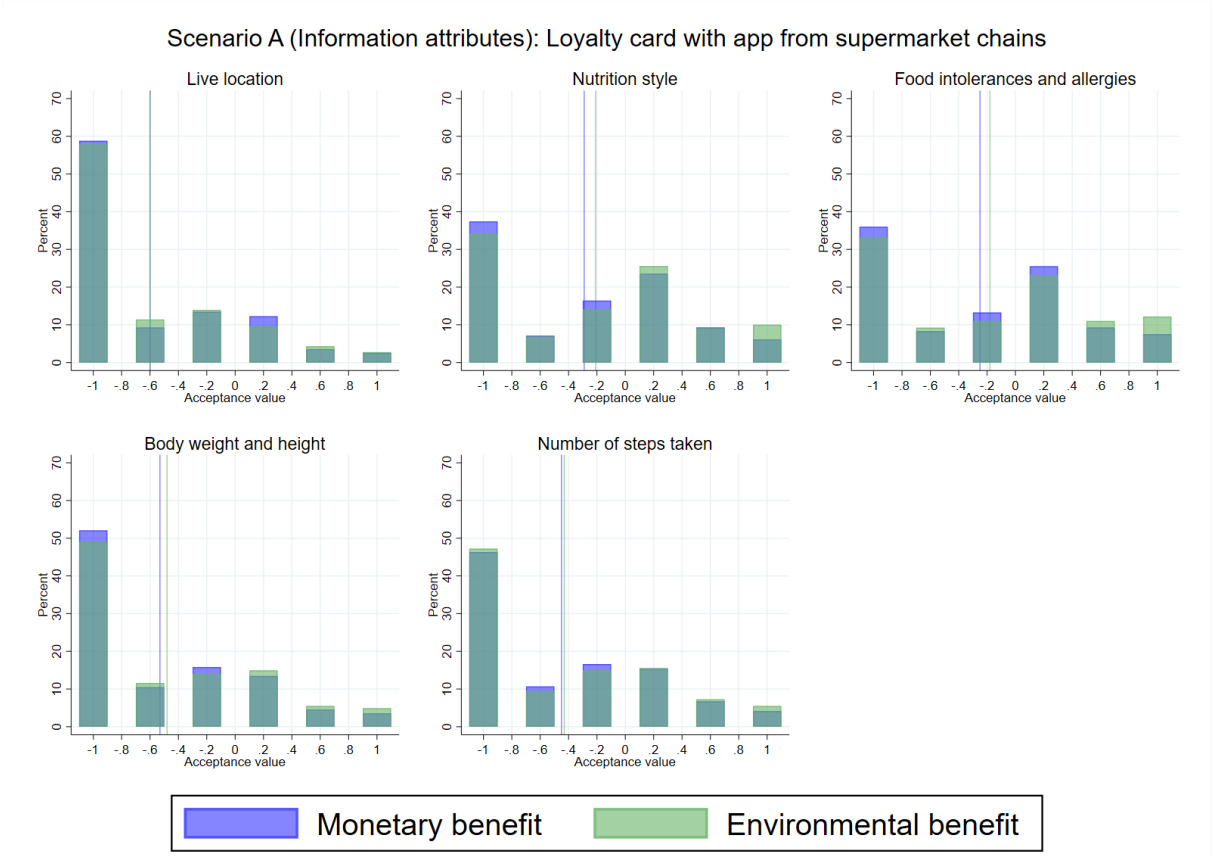


Figure 5.7: Histograms showing the percentage distribution of acceptance values for optional information attributes in scenario A by treatment.

Notes. Vertical lines indicate the mean values in T1 (blue) and T2 (green).

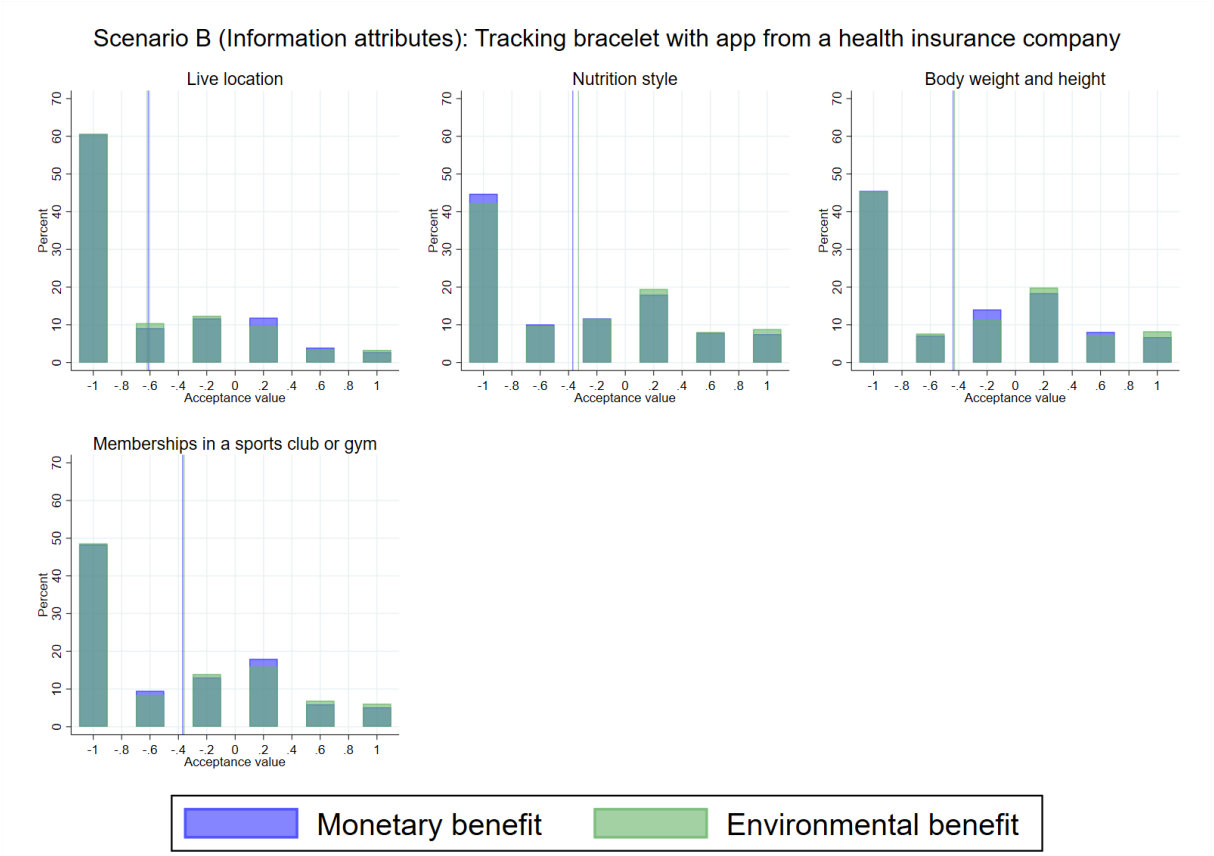


Figure 5.8: Histograms showing the percentage distribution of acceptance values for optional information attributes in scenario B by treatment.

Notes. Vertical lines indicate the mean values in T1 (blue) and T2 (green).

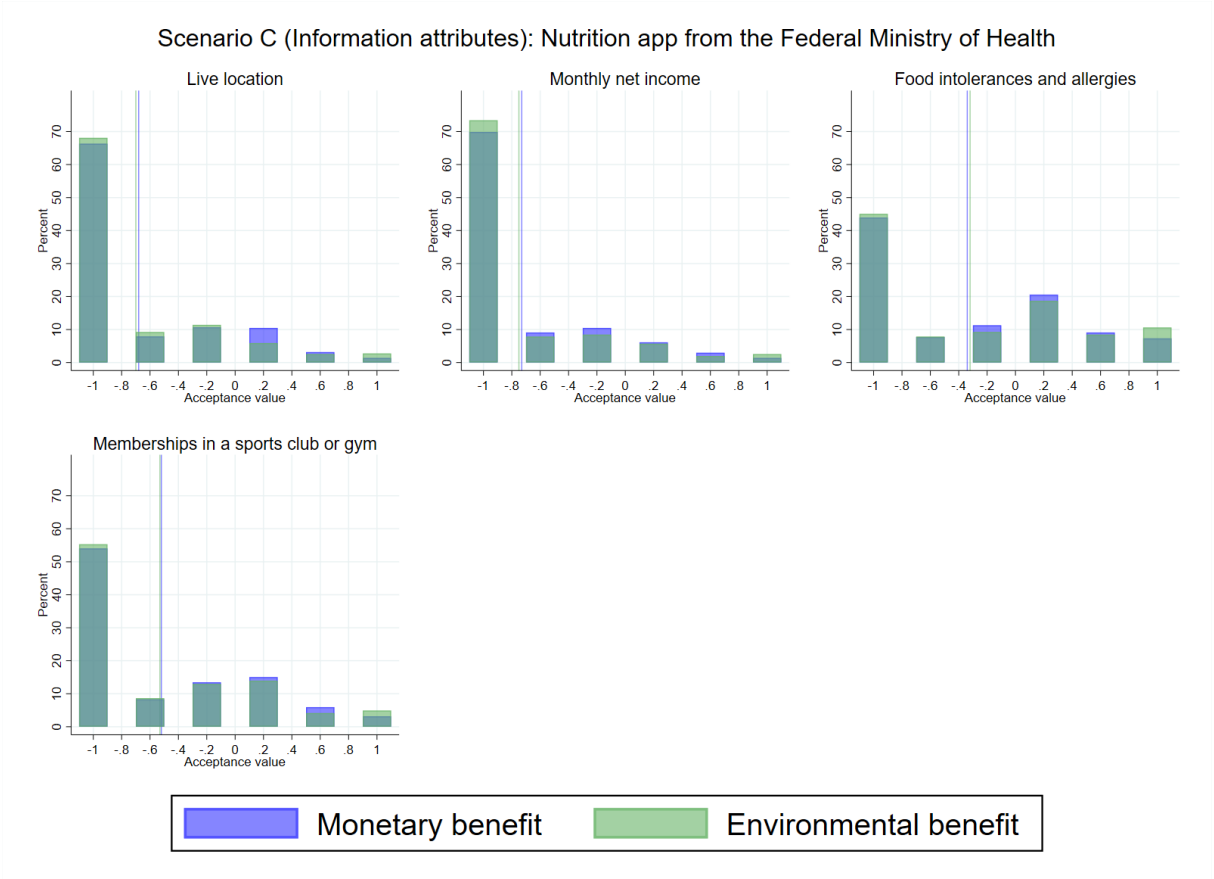


Figure 5.9: Histograms showing the percentage distribution of acceptance values for optional information attributes in scenario C by treatment.

Notes. Vertical lines indicate the mean values in T1 (blue) and T2 (green).

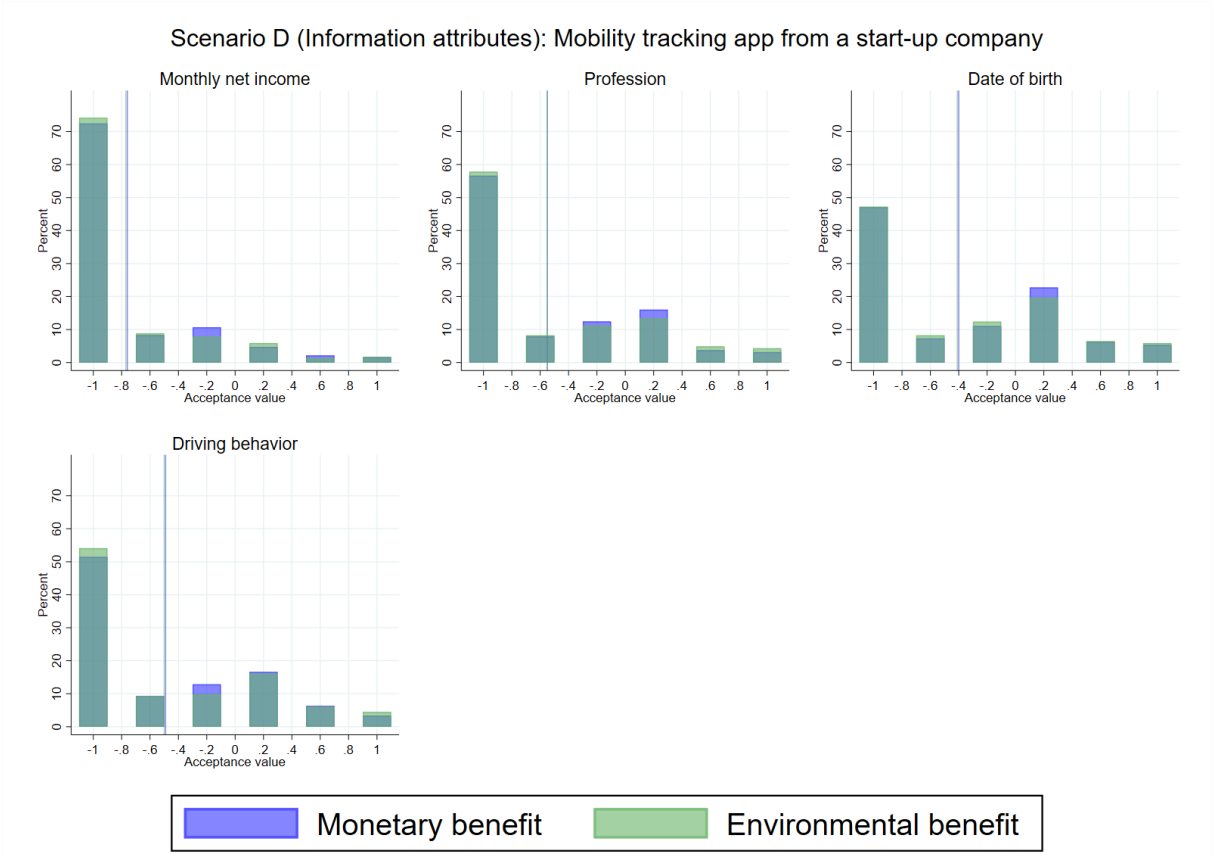


Figure 5.10: Histograms showing the percentage distribution of acceptance values for optional information attributes in scenario D by treatment.

Notes. Vertical lines indicate the mean values in T1 (blue) and T2 (green).

Figure 5.11: Histograms showing the percentage distribution of acceptance values for optional information attributes in scenario E by treatment.

Notes. Vertical lines indicate the mean values in T1 (blue) and T2 (green).

5.6.4 D – Pairwise tests on differences between optional recipients or optional information attributes

Scenario A Loyalty card	Recipient	Household members	Employer	Federal ministries	Market research comp.	German food prod.	American food prod.	Chinese food prod.
Recipient	Acc. value	-0.21	-0.67	-0.39	-0.18	-0.26	-0.64	-0.70
Employer	-0.67	***						
Federal ministries	-0.39	***	***					
Market research comp.	-0.18	not sig.	***	***				
German food producers	-0.26	**	***	***	***			
American food prod.	-0.64	***	not sig.	***	***	***		
Chinese food prod.	-0.70	***	not sig.	***	***	***	***	
Health insurance comp.	-0.48	***	***	***	***	***	***	***

Table 5.11: Wilcoxon signed rank tests for significant differences between the acceptance values for optional recipients in scenario A.

Notes. Average acceptance values for combined treatments are shown in the right column next to each recipient and in the second row. * $p < 0.01$, ** $p < 0.005$, *** $p < 0.001$.

Scenario B Tracking bracelet	Recipient	Household members	Employer	Federal ministries	Market research comp.
Recipient	Acc. value	-0.31	-0.70	-0.46	-0.38
Employer	-0.70	***			
Federal ministries	-0.46	***	***		
Market research comp.	-0.38	***	***	***	
German sports equip. prod.	-0.54	***	***	***	***

Table 5.12: Wilcoxon signed rank tests for significant differences between the acceptance values for optional recipients in scenario B.

Notes. Average acceptance values for combined treatments are shown in the right column next to each recipient and in the second row. * $p < 0.01$, ** $p < 0.005$, *** $p < 0.001$.

Scenario C Nutrition app	Recipient	Household members	Employer	Market research comp.	German food prod.
Recipient	Acc. value	-0.40	-0.75	-0.43	-0.52
Employer	-0.75	***			
Market research comp.	-0.43	not sig.	***		
German food prod.	-0.52	***	***	***	
German Society for Nutrition	-0.40	not sig.	***	**	***

Table 5.13: Wilcoxon signed rank tests for significant differences between the acceptance values for optional recipients in scenario C.

Notes. Average acceptance values for combined treatments are shown in the right column next to each recipient and in the second row. * $p < 0.01$, ** $p < 0.005$, *** $p < 0.001$.

Scenario D Mobility tracking	Recipient	Household members	Employer	Federal ministries	Public transport comp.	Car insurance comp.
Recipient	Acc. value	-0.41	-0.76	-0.52	-0.51	-0.66
Employer	-0.76	***				
Federal ministries	-0.52	***	***			
Public transport comp.	-0.51	***	***	not sig.		
Car insurance comp.	-0.66	***	***	***	***	
City or municip.	-0.58	***	***	***	***	***

Table 5.14: Wilcoxon signed rank tests for significant differences between the acceptance values for optional recipients in scenario D.

Notes. Average acceptance values for combined treatments are shown in the right column next to each recipient and in the second row. * $p < 0.01$, ** $p < 0.005$, *** $p < 0.001$.

Scenario E Smart meter	Recipient	Household members	Employer	Federal ministries	German elect. prod.	Neigh-bors
Recipient	Acc. value	-0.20	-0.79	-0.51	-0.49	-0.78
Employer	-0.79	***				
Federal ministries	-0.51	***	***			
German elect. prod.	-0.49	***	***	not sig.		
Neighbors	-0.78	***	not sig.	***	***	
City or municip.	-0.49	***	***	not sig.	not sig.	***

Table 5.15: Wilcoxon signed rank tests for significant differences between the acceptance values for optional recipients in scenario E.

Notes. Average acceptance values for combined treatments are shown in the right column next to each recipient and in the second row. * $p < 0.01$, ** $p < 0.005$, *** $p < 0.001$.

Scenario A Loyalty card	Information attribute	Food intolerances and allergies	Live location	Nutrition style	Body weight and height
Information attribute	Acc. value	-0.22	-0.60	-0.25	-0.50
Live location	-0.60	***			
Nutrition style	-0.25	not sig.	***		
Body weight and height	-0.50	***	***	***	
Number of steps taken	-0.44	***	***	***	***

Table 5.16: Wilcoxon signed rank tests for significant differences between the acceptance values for optional information attributes in scenario A.

Notes. Average acceptance values for combined treatments are shown in the right column next to each recipient and in the second row. * $p < 0.01$, ** $p < 0.005$, *** $p < 0.001$.

Scenario B Fitness bracelet	Information attribute	Memberships in a sports club or gym	Live location	Nutrition style
Information attribute	Acc. value	-0.44	-0.61	-0.35
Live location	-0.61	***		
Nutrition style	-0.35	***	***	
Body weight and height	-0.37	***	***	not sig.

Table 5.17: Wilcoxon signed rank tests for significant differences between the acceptance values for optional information attributes in scenario B.

Notes. Average acceptance values for combined treatments are shown in the right column next to each recipient and in the second row. * $p < 0.01$, ** $p < 0.005$, *** $p < 0.001$.

Scenario C Nutrition app	Information attribute	Memberships in a sports club or gym	Live location	Food intol. and allergies
Information attribute	Acc. value	-0.52	-0.69	-0.33
Live location	-0.69	***		
Food intol. and allergies	-0.33	***	***	
Monthly net income	-0.74	***	***	***

Table 5.18: Wilcoxon signed rank tests for significant differences between the acceptance values for optional information attributes in scenario C.

Notes. Average acceptance values for combined treatments are shown in the right column next to each recipient and in the second row. * $p < 0.01$, ** $p < 0.005$, *** $p < 0.001$.

Scenario D Mobility tracking	Information attribute	Driving behavior	Profession	Date of birth
Information attribute	Acc. value	-0.50	-0.55	-0.40
Profession	-0.55	**		
Date of birth	-0.40	***	***	
Monthly net income	-0.77	***	***	***

Table 5.19: Wilcoxon signed rank tests for significant differences between the acceptance values for optional information attributes in scenario D.

Notes. Average acceptance values for combined treatments are shown in the right column next to each recipient and in the second row. * $p < 0.01$, ** $p < 0.005$, *** $p < 0.001$.

Scenario E Smart meter	Information attribute	Live location	Profession	Time and dur. of use of ind. power sources
Information attribute	Acc. value	-0.69	-0.65	-0.25
Profession	-0.65	**		
Time and dur. of use of ind. power sources	-0.25	***	***	
Monthly net income	-0.78	***	***	***

Table 5.20: Wilcoxon signed rank tests for significant differences between the acceptance values for optional information attributes in scenario E.

Notes. Average acceptance values for combined treatments are shown in the right column next to each recipient and in the second row. * $p < 0.01$, ** $p < 0.005$, *** $p < 0.001$.

5.6.5 E – Scales on privacy concerns, risk, and GREEN consumption values

Item	Question	Mean (sd)
PC1	I feel uncomfortable when my online behaviors are tracked without permission.	5.63 (1.54)
PC2	I am concerned about misuse of my online behaviors.	5.27 (1.46)
PC3 (R)	It does not bother me to receive too much advertising material through tracking of my online behaviors.	5.20 (1.81)
PC4	I fear that my online behavior information may not be safe while stored.	5.20 (1.44)
PC5 (R)	I do not believe that my online behavioral data is often misused.	4.63 (1.68)
PC6 (R)	I do not think companies share my online behavioral data without permission.	4.76 (1.90)
All six		5.12 (1.03)
Answer: (1) Strongly disagree; (2) Somewhat disagree; (3) Rather disagree; (4) Neither nor; (5) Rather agree; (6) Somewhat agree; (7) Strongly agree		

Table 5.21: Privacy concern scale (Ham, 2017).

Notes. Means and standard deviations (in parentheses). (R) denotes reverse items.

Question	Mean (sd)
How do you see yourself: are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please tick a box on the scale, where the value 0 means: 'not at all willing to take risks' and the value 10 means: 'very willing to take risks'.	4.33 (2.52)

Table 5.22: Risk attitude question (Dohmen et al., 2011).

Notes. Mean and standard deviation (in parentheses).

Item	Question	Mean (sd)
GC1	It is important to me that the products I use do not harm the environment.	5.09 (1.33)
GC2 (R)	I do not consider the potential environmental impact of my actions when making many of my decisions.	4.24 (1.61)
GC3	My purchase habits are affected by my concern for our environment.	4.22 (1.63)
GC4 (R)	I am not concerned about wasting the resources of our planet.	5.31 (1.75)
GC5	I would describe myself as environmentally responsible.	4.92 (1.26)
GC6 (R)	I am not willing to be inconvenienced in order to take actions that are more environmentally friendly.	4.45 (1.67)
All six		4.71 (0.99)
Answer: (1) Strongly disagree; (2) Somewhat disagree; (3) Rather disagree; (4) Neither nor; (5) Rather agree; (6) Somewhat agree; (7) Strongly agree		

Table 5.23: GREEN consumption scale (Haws et al., 2014).

Notes. Means and standard deviations (in parentheses). (R) denotes reverse items.

6. Recycling behavior of private households: an empirical investigation of individual preferences in a club good experiment

This chapter has been published as:

Menges, R., Cloos, J., Greiff, M., Wehrle, J., Goldmann, D., & Rabe, L. (2021). Recycling behavior of private households: an empirical investigation of individual preferences in a club good experiment. *Clean Technologies and Environmental Policy*, *23*, 843-856.

This paper can be downloaded via

https://link.springer.com/article/10.1007/s10098-020-01929-5

Abstract: While recycling helps to limit the use of primary resources, it also requires considerable technological investments in regional circular flow systems. The effectiveness of recycling systems, however, also depends on household behavior. Therefore, current research increasingly focuses on behavioral and psychological theories of altruism, moral behavior, and social preferences. From an economic perspective, recycling systems can be understood as public goods with contributions resulting in positive externalities. In this context, the literature shows that recycling behavior highly depends on the perception of how others behave. In neutrally framed public good experiments, contributions tend to increase when alternative public goods are offered and group identity is generated. We aim to contribute to this discussion by observing household behavior concerning recycling opportunities in controlled settings. For this purpose, we study a laboratory experiment in which individuals contribute to recycling systems: At first, only one public recycling system (public good) is offered. After dividing societies into two clubs, "high" and "low" according to their environmental attitudes, excludable club systems (club goods) are added as alternative recycling options for each club. The results of our pilot experiment show that adding a more exclusive recycling club option increases individual contributions to recycling compared with a pure public good framework. However, this increase in cooperation is only significant for those clubs where members with higher environmental attitudes are pooled.

Keywords: experimental economics, recycling, recycling policy, external effects, public goods, club goods

6.1 Introduction

In recent years, the intensification of industrial recycling activities and the improvement of the associated circular economy networks have become an essential area of environmental policy. Recycling does not only help to reduce the environmental costs caused by waste, but can also be used specifically by industrial policy to reduce certain shortages of raw materials, as in the case of rare elements that are used in mobile phones. In addition, recycling can be used to reduce the carbon footprint of consumption activities, helping to mitigate the required space and emissions from landfills all around the world. In this paper, we use the broader definition of the term "recycling" compared with the narrower one in the European waste legislation. Accordingly, the term comprises all means of how to feed waste from components and materials back into the production and consumption cycle, including all necessary collection, transportation, and treatment steps. Therefore, recycling, in the sense addressed in this paper, includes product and component reuse, use after remanufacturing and second life applications, material and feedstock recycling, as well as energy recovery. The term "discarding" as used in this paper characterizes the least efficient, yet legal, orderly way to get rid of waste in terms of environmental and resource efficiency aspects. A third option not addressed in this paper comprises the more or less illegal ways of littering. The term (legal) "disposal" incorporates all means of how to bundle and transfer public waste streams to legal collection systems and pass these on to elaborated recycling structures, to be transferred down to low-level discarding systems.

From an economic point of view, recycling is not cost free and does not represent a certain kind of a perpetual motion machine which allows to overcome the restrictions of the laws of thermodynamics. However, in many cases, it is cheaper and more efficient to recycle the raw materials contained in the products consumed, instead of disposing them in the waste. Even though raising recycling rates is a task of environmental policy, the effectiveness of recycling systems depends on cooperative household behavior. Transferring specific waste streams to elaborated collection and recycling structures may increase opportunity costs for private households by raising the personal effort of informing, reflecting, sorting, and depositing in comparison with collective disposal at the most convenient point.

Economic literature highlights positive externalities in the case of recycling. Households that voluntarily participate in recycling systems also contribute to the internalization of external effects associated with conventional waste management. For at least four decades, recycling-

related behavior of private households has been investigated in a large number of empirical studies (such as Kinnaman, 2009; Beede and Bloom, 1995). These studies, for instance, address the question of which regulatory conditions or which social norms and motives have a positive effect on recycling behavior. Although several papers on recycling behavior marginally consider elements of strategic interaction and mutual influence within peer groups, most of them lack explicit game-theoretical elements where individuals voluntarily contribute to public goods. In this paper, we present an incentivized economic experiment where subjects adopt the role of private households and have various options for recycling and waste disposal. The underlying model treats recycling as a social dilemma situation. Within a society, each recycling decision has positive external effects on all other households. While it is beneficial for the society as a whole if every household is cooperative and involved in recycling as fully as possible, each single household has an incentive to deviate from that kind of cooperation. Moreover, we give households the opportunity not only to participate in large public recycling systems, but also to contribute to smaller, more exclusive recycling clubs whose members, to a large degree, also share specific environmental attitudes. The contribution of our article is, therefore, the examination of the following two research questions: First, we ask whether the expansion of recycling options leads to an increase in household contributions (at all); second, we investigate whether the different composition of the clubs affects the decisions of their members.

Our paper is organized as follows: Section 6.2 presents a short overview of the empirical literature on the recycling behavior of private households. In section 6.3, we discuss the suitability of the club good model for an experimental investigation of recycling behavior. Section 6.4 entails a description of the experimental design we employ and is followed by a presentation of the results of our study in section 6.5. The last section offers a summary of the main observations and derives some conclusions with respect to the design of further research.

6.2 Recycling behavior of private households

6.2.1 State of the theoretical and empirical literature

Environmental economics treats recycling as a classic case of market failure. In this respect, there are external effects that differ from public goods only in terms of their aggregation to the optimal quantity. The generation of waste can be interpreted as a by-product of private consumption and production activities inducing costs for the whole society, which are not incorporated in market prices. Over the last three decades, economists have investigated

different institutional arrangements of how to organize recycling systems on governmental or municipal levels and emphasized the external cost argument for legitimizing such a type of market intervention. Empirical estimates of external costs of the disposal of different types of solid waste are used for analyzing costs and benefits of certain recycling systems (Kinnaman, 2009; Beede and Bloom, 1995). According to environmental economics textbooks, recycling problems, when regarded as problems of external effects, can be solved by applying environmental policy instruments such as deposit systems. Suboptimal recycling rates can even be raised to optimal levels if the "true" level of external costs is known (Tietenberg and Lewis, 2018). However, taking into account that perfect information of environmental regulators is not realistic to assume, and that perfect knowledge about consumer behavior is not available, things become more complicated. Environmental policies directed at increasing recycling rates in the economy are strongly reliant on cooperative behavior of private households as private waste management activities cannot be monitored or regulated directly without imposing massive regulations on all citizens.

There is a bulk of empirical literature on household behavior in the recycling case which has grown over the last decades. In a recent survey, Briguglio (2016) analyzes empirical papers that investigate conditions and regulatory approaches that potentially stimulate household cooperation in the case of recycling. Briguglio shows that, more recently, the focus of empirical literature has shifted from analyzing theoretical models of waste production (including illegal disposal and recycling) to behavioral and psychological theories of altruism, moral behavior, and social preferences.

Trudel (2019) suggests that the psychological factors that influence recycling behavior can be divided into four areas, namely cognitive barriers, the self, social influence, and product characteristics. Certain product characteristics, for example, might give rise to behavioral effects in the case of recycling. It can be shown that distorting products from their original form, e.g., by cutting them up, leads to less recycling activities as compared to recycled products that have kept their original design (Trudel et al. 2016). Among others, Schultz (1999) and Barr (2007) investigate recycling behavior from the perspective of social psychology. Schultz (1999) shows that voluntary participation and individual contributions in recycling programs can be extended if the social environment of the individual household allows for group feedbacks which interact with personal norms. The results of Barr (2007) suggest that a certain mix of subjective environmental attitudes, situational characteristics, and psychological

factors can be used to explain recycling behavior. However, in contrast to household behavior in the field of waste reduction or private reuse of materials, which can be explained merely by subjective knowledge or concern-based factors, recycling behavior has to be interpreted as a "highly normative behavior." This conclusion is also supported by a study of Cecere et al. (2014) which investigates the interplay of extrinsic and intrinsic motivation in the case of private households' waste management activities. Cecere et al. distinguish between waste reduction and recycling, and show that the behavior of households in reducing their waste can hardly be explained by social or group-oriented motives as it is not an observable action. Rather, it can be linked to purely altruistic motives. Against this backdrop, extrinsic motivation associated with social norms, imitative behaviors, and peer pressure are important for recycling activities, because they are visible in society and are, for example, open to reciprocity considerations. Indirectly, they support the standard hypothesis mentioned above, namely that recycling activities which are not driven by intrinsic motivation can be incentivized by economic instruments (such as taxes, levies, or deposit systems). Interestingly, there are several papers such as Kaoursakis and Birol (2008) which state that while there is a general public acceptance for introducing incentive-oriented instruments, households prefer the introduction of deposit refund schemes rather than unit-pricing programs or pay-as-you-throw schemes.

The role of social norms and warm-glow elements of altruistic behavior in the case of recycling is also highlighted by Abbott et al. (2013). They conclude that economic incentives should not be used as a substitute but rather as a complement for intrinsic values. Social norms and group-induced effects play a greater role than subjective environmental attitudes. At the same time, their results refute an influence of warm-glow motifs. Politicians might therefore rely on social norms rather than influence behavior directly. These results also underline the importance of decentralized, supply-side activities such as kerbside collection, which makes the participation of individual households more visible to their neighborhood.

6.2.2 Research gap

Other studies such as Brekke et al. (2010) investigate strategic interactions of individuals and highlight the public good feature of individual recycling decisions. They start by questioning why individual contributions to recycling systems are increasing in contributions of other individuals, as reported in several studies. Similar to other cases of environmentally friendly behavior of consumers, such as green electricity (Menges et al., 2005), Brekke et al. (2010) ask whether contributions of other individuals affect individual donations as complements or as

substitutes. The authors contrast impure altruism to individuals' duty orientation when contributing to glass recycling systems. The hypothesis of duty orientation states that individuals driven by external norms and duties interpret behavior of other people as an indication of their own responsibility to participate. Brekke et al. (2010) point out that the concept of responsibility feelings, which are somewhat unusual in economics, is related to the behavioral concept of reciprocal preferences, which is well documented in experimental economics (see Fehr and Gächter, 2000; Hoffman et al., 1998). On the other hand, impure altruism and warm-glow motivation are described by pure pleasure motivation and less by social interaction. The empirical results of the study on recycling behavior reported by Brekke et al. (2010) point to strong social interaction effects which can be linked to the attribution of responsibility. Households who firmly believe that their recycling attitudes are shared within their group derive a greater sense of responsibility from these attitudes. According to Brekke et al. (2010), this observation is in line with the hypothesis that duty and responsibility cannot be interpreted as exogenous factors when analyzing preferences. Households derive conclusions about their own responsibility by observing other households. Although such a type of responsibility constitutes a burden, households are willing to accept this burden and to engage in recycling if the duty is indisputable. Brekke et al. (2010) conclude that policy approaches such as public campaigns for recycling have merely a stimulating effect if they change the perception of what others are doing.

Further authors describe the connection between recycling and the theory of public goods. Rompf et al. (2017) point out that institutional trust can counteract the inhibitory effect of individual costs on cooperation in a collective action dilemma, using recycling as an example. Huhtala (2009) finds a negative income effect on the willingness to pay for recycling, but a positive income effect on the willingness to pay for more convenient incineration of waste. This raises distributional issues, since poorer households make comparatively greater efforts to behave in an environmentally sound manner.

Another approach worth mentioning is the method of "Identity Economics," which was presented by Akerlof and Kranton (2000) and whose relevance was recently underlined by economists such as Foss (2019) and Collier (2019). Akerlof and Kranton (2000) incorporate the factor identity into the economic behavioral model. They assume that societies consist of different social categories that share certain norms. Using a simple game-theoretic model, they show that members of social categories are inclined to reduce deviations from these norms and

adapt their behavior accordingly. From an economic perspective, further investigation of these considerations, e.g., with respect to recycling, seems promising.

Summed up, numerous empirical studies reveal factors influencing recycling behavior. Although many highlight social interaction phenomena of household behavior when people contribute to public goods on a voluntary basis, strategic elements of individual behavior in the case of recycling situations have not been considered in the literature so far. Therefore, we make a first attempt to investigate interaction effects in different institutional settings to determine which role can be ascribed to individual environmental attitudes. The experimental design and the hypotheses which are tested are drawn from the model which is presented in the next section.

6.3 Model

6.3.1 Recycling in a public good framework

The starting point of our model is as follows: Voluntary participation and individual contributions to a recycling system entail positive externalities for the whole society. This structure allows for modeling a public goods game. A recycling system can, by definition, be treated as a public good: Its output (such as the avoidance of garbage and the protection of scarce resources) is represented by the sum of voluntary individual contributions. However, its benefits are also accessible to members of the society who have not participated in the provision (recycling process). Given certain assumptions, this formal model can be solved with respect to individual utility maximizing behavior (Nash equilibrium) and its social optimum (welfare maximization). In its simplest version, the model can be represented as follows (Ledyard, 1995): Assume that a society consists of N households. Each single household i has an endowment of m_i which is used for consumption. We assume that consumption generates waste one by one. Hence, each household has to derive a decision on how to manage m_i units of waste. The household can contribute g_i $(0 \leq g_i \leq m_i)$ units of its waste to a public recycling system. The payoff u_i each household realizes from this decision is determined by the following equation:

$$u_i = (m_i - g_i) + a \sum_{j=1}^{N} g_j \qquad (1)$$

The opportunity cost of contributing to the recycling system (e.g., due to inconvenience) is normalized to one per unit. Alternatively, waste can be dumped to trash cans (which is usually the case for mixed residual waste) without any direct or immediate cost. The benefit realized

per unit of g_i is set to a. Hence, the marginal per capita return (MPCR) of contributing to the public recycling system is a (with $0 < a < 1 < Na$) and represents the gap between individual utility maximization ($g_i = 0$) and the Pareto-efficient social optimum ($g_i = m_i$) which is driven by positive externalities. Hence, the question arises of how to close this gap and to overcome such a kind of market failure.

The empirical literature reported in the previous section also discusses the issue of varying the supply-side of public recycling systems in several dimensions, such as its opportunity cost structure (e.g., convenience effects of kerbside versus non-kerbside systems) or the number of resources and materials collected. There are several papers (such as Abbott et al., 2017; Kaoursakis and Birol, 2008) that investigate how households' willingness to contribute reacts to such supply-side variations.

6.3.2 Simultaneous provision of public goods and club goods

In our model, we also vary the supply-side of recycling systems by introducing an additional feature. Public recycling systems can also be organized as club systems. The most important difference between a public good and a club good lies in the exclusion principle: While the benefits jointly generated in club goods are accessible to all members of the club, they are not accessible to all other members of society. From an economic perspective, this exclusion principle gives rise to an interesting point: Although the same problem of cooperation arises within the boundaries of the club as with public goods (e.g., free riding), the members of the club can better protect themselves from noncooperative behavior by virtue of the exclusionary principle. This potential feature of club goods is related to the results of Brekke et al. (2010) obtained from a survey of households' glass recycling activities in Norway. They show that households that are aware of sharing recycling attitudes within their peer group derive a greater sense of responsibility and contribute more to the recycling system. This observation is in line with the general hypothesis of Mancur Olson (1965): Commonly shared goals face organizational problems in large groups when individuals perceive only limited relevance of their actions. Therefore, adding elements of excludability and selective "private" benefits become an important strategy to overcome this social dilemma.

The importance of an exclusion mechanism of clubs is also emphasized by Cornes and Sandler (1996). Therefore, it can be necessary to exclude individuals who do not invest in the club's purpose in a satisfactorily manner in order to prevent free riding. However, the

introduction of the club good with its theoretical exclusion possibilities does not represent a change in the material incentive structure in this model. Firstly, there is still the possibility of investing in the public good and, secondly, the attraction of free-riding in the club good context remains in the same form. The introduction of the club good hence represents a change in the social context of the interaction of individuals within a predetermined society. We are thus taking up a central idea in the field of the so-called Identity Economics (Akerlof and Kranton, 2000). According to this theory, it is not questioned whether the individual behavior can be explained by cost/benefit calculations. However, the influence of these motives can be changed by the social context in which individuals act and see themselves as part of a group. It is a well-known result in the experimental literature that individuals tend to treat their counterparts more kindly if they themselves are treated kindly by them (Rabin, 1993).

The empirical literature on club goods is somewhat limited in experimental economics. Although individual decisions in such types of public good experiments are mostly framed as charitable giving, results can also be interpreted in a general manner or with respect to the recycling case. Chakravarty and Fonseca (2017) show that cooperation and the overall contribution of individuals substantially increase when they are offered an additional group-specific investment opportunity. However, these efficiency gains take place only when the expected financial returns to the club good are at least equivalent to the public good. An experiment by Blackwell and McKee (2003) identifies two drivers of individual contributions to club goods and public goods when investment opportunities are offered simultaneously. Contributions to the group-specific good are characterized by reciprocity as they increase with past contributions of other group members. The allocation of spending for the club good and the public good is affected by the respective MPCR. The standard result of experimental economics posits that some level of cooperation in public good environments still exists but decreases from period to period. Quite interestingly, it can be rejected for the club good investment as in Blackwell and McKee (2003) where it does not decrease over time. An experiment by Cherry and Dickinson (2008) suggests that there is a general tendency that cooperation and total contributions increase when individuals are offered multiple sets of alternative public goods. Chen and Li (2009) show that the generation of group identities in such kind of experiments can have a large impact on social preferences, for instance by reducing envy effects. Several studies investigate constitutional effects of group identities, such as natural identities (Chen et al., 2014) or gender (Brown-Kruse and Hummels, 1993), which might give rise to conditional, group-specific altruism.

The public good model (1) can easily be supplemented with an additional club good opportunity. Assume that the society can be divided into two distinct subgroups with an equal number of households $N/2$ and that each household i also has the opportunity to invest k_i $(0 \leq k_i \leq m_i)$ units of their waste to a group-specific recycling system. The return of this investment and its MPCR is represented by b (with $0 < b < 1 < (N/2)\, b$). The payoff of the waste management decision of household i can be expressed as:

$$u_i = (m_i - g_i - k_i) + a \sum_{j=1}^{N} g_j + b \sum_{l=1}^{N/2} k_l \tag{2}$$

As in (1), in line with economic standard theory, utility maximizing behavior by choosing g_i and k_i predicts that $g_i = k_i = 0$ (Nash equilibrium). On the other hand, the Pareto-optimal solution recommends maximum contributions of all individuals. If $a = b$, the welfare solution is indifferent to different allocations among g_i and k_i, as long as $g_i + k_i = m_i$. However, if both goods have the same MPCR ($a = b < 1$), then different group sizes might induce a different return on the investment if and only if specific fractions of other members are expected to contribute as well. While some authors argue in line with Olson (1965) and assume decreasing average contributions for an increasing N in public good games, Ledyard (1995) reports several studies on group size effects, indicating that the public good is potentially more attractive as it has a larger number of potential contributors. For large groups with ten players and small groups with four players, Isaac and Walker (1988) show that for a small MPCR, average contributions to the public good are higher in the large groups and that for a large MPCR there is no clear effect of group size. In order to compensate for this group size effect, we choose an approach suggested by Chakravarty and Fonseca (2017) and Blackwell and McKee (2003) and normalize the MPCR as follows: $a\, N = b\, (N/2)$ which is equal to $2a = b$. This approach is also in line with the concept of the effective average per capita ratio (APCR) which is $(aN)/N$ for the public good and $b(N/2)/N$ for the club good. Whereas the MPCR to the individual is a in the case of the public good and b in the case of the club good, the APCR divides the total return of investing one unit by all members of the society. Although, from a marginal perspective, it is clear that this return of the club good is actually not distributed among all members of society, it is established as useful behavioral concept for comparing the relative payoffs with the group and the public good (Blackwell and McKee, 2003). Note that if $aN = b(N/2)$, the Pareto-efficient welfare optimum can include different allocations among g_i and k_i, as long as $g_i + k_i = m_i$.

Applying this club good structure (2) and its intergroup competition to the recycling case also implies introducing an element of institutional competition of different recycling systems on the supply-side. As an alternative to free waste disposal, each household has two recycling systems at its disposal: A system that can be used for the economy as a whole and a certain closed-shop recycling system, where access and benefits are exclusive to members. In reality, private households, for example, have access to municipal recycling yards and collection points which offer their services exclusively to citizens of the municipality on the basis of public fee schedules. An example of such a disparate benefit could be that the improved recycling efforts of club members result in a cost reduction of the public fee structure due to a better utilization of recycling capacities. At the same time, households can often hand over their waste to other commercial traders, who then collect it throughout voluntary collections. In addition, recycling also takes place in the course of take-back obligations of commercial trade.

In our experiment, two types of recycling systems are modeled as the public good and the club good. In both recycling systems, monetary incentives are identical, e.g., contributing one unit to the public good or the club good reduces the own payoff by one unit but increases social welfare (measured as the sum of *N* payoffs) by $2a = b$.

6.4 Experimental design

6.4.1 Treatment structure and procedure

Since this recycling experiment forms part of a larger interdisciplinary research project on sustainable household behavior, it serves to identify key points for longer-term future investigations. The pilot reported in this paper consists of two treatments: In order to test for the relevance of the club good structure, the first treatment (treatment 1) refers to the public good model (1) and the second treatment (treatment 2) is based on the club good model (2). The experiment was conducted as a classroom experiment with 36 undergraduate students during an economics course at the Clausthal University of Technology in Germany on 16 July 2019. Table 6.1 provides an overview of the experimental design.

Treatment	Treatment 1	Treatment 2
Public good MPCR	0.3	0.3
Club good MPCR	–	0.6
Endowment (waste units)	10	10
Number of rounds	10	10
Number of societies	6	6
Subjects	36	36

Table 6.1: Experimental design.

For the experiment, we used classEx software (Giamattei and Lambsdorff, 2019) and a within-subjects design. This means that participants are exposed to both treatments, one after the other, and each participant serves as his or her own baseline. A main advantage of this design is that it may reduce the variance of unobserved components, thereby increasing the precision of the estimated average treatment effect. As compared to other design options, fewer subjects have to be recruited (Czibor et al., 2019; List et al., 2011). Moreover, the nonparametric tests used in the following sections work well with our sample size. We designed a repeated game over ten rounds per treatment. This approach enables us to take learning effects into account (Ledyard, 1995) and corresponds to reality because waste disposal decisions have to be made repeatedly. However, due to organizational limitations of the software, we are not able to control for sequence effects. For further experiments, we will switch to a more appropriate between-subjects design. The experiment took 45 min, and the exact procedure can be summarized as follows:

- Subjects were informed that in the experiment, they would adopt the perspective of households which had to decide about their waste disposal. They were also informed that the task was to decide on the use of an initial endowment. A monetary amount reflecting the results of these decisions would be paid out in cash on the basis of a lottery at the end of the experiment (see later section).
- After connecting mobile devices to the server, participants underwent a questionnaire with respect to certain environmental attitudes which were used for later group separation (see later section).
- Subjects were told that they had been randomly assigned to a society of six households altogether and that the composition of this society would remain unchanged throughout the experiment. They were also informed that they did not know the other households with which they formed a society and that there was no way to communicate with them (stranger design).

- Thereupon, printed instructions for treatment 1 were handed out, and—after they were read by the participants - the corresponding game was played.
- After finishing the final round of treatment 1, participants were informed about the formation of the clubs within their society. They were assigned to "club 1" or "club 2," each consisting of three households. Then, a so-called slider-game was played where the different clubs competed against each other (identity reinforcement, see later section).
- Once again, instructions were handed out, and the second treatment (treatment 2) was conducted.
- The experiment ended after collecting some demographic data.
- Finally, participants received their payment according to the lottery mechanism (incentive mechanism, see later section).

6.4.2 Instructions and framing

By reading the instructions for treatment 1, participants are introduced to the following situation: Every participant adopts the role of a single household which is part of a society of six households. Each household of this society is endowed with ten tokens in each round which are used for consumption. When consuming, waste is generated and has to be disposed. Two options are offered: Ten units of waste can be disposed either as conventional household waste with no additional costs or in a public recycling system. For every unit of waste which is brought to this recycling system, one token has to be paid. Due to positive externalities of recycling, each recycled unit generates a benefit of 0.3 tokens for all six members of the society, as expressed in (1). The decision has to be made repeatedly in ten consecutive rounds. After each round, each player is informed about the investments and payoffs of all members of his or her society.

The instructions for the subsequent treatment 2 pick up from the previous framing of recycling opportunities but introduce an additional disposal option as suggested in Eq. (2): Alongside with the possibilities known from treatment 1, the ten units of waste can also be recycled in a system that is only accessible for its three club members. Again, costs to dispose waste in this system are one unit per token, while the benefit of 0.6 tokens per recycled unit is reserved for the three club members. Participants were aware that both recycling options are similarly attractive with respect to their social returns, because $6 \cdot 0.3 = 3 \cdot 0.6$. As before, this decision has to be repeated ten times. The structure of the decision screen is presented in the Appendix.

6.4.3 Introducing and enforcing group identity

Adapting the club good model (2) to the recycling case in the context of a laboratory experiment involves the question of how to design a meaningful grouping variable which can be used to split a society of N subjects into nonoverlapping groups. The experimental design of the few club good experiments reported in the literature follows the so-called minimal group paradigm. In these approaches, subjects are grouped by asking them to solve simple tasks such as evaluating paintings (Chakravarty and Fonseca, 2017; Chen and Li, 2009) or by exogenously assigning subjects to different colors (Blackwell and McKee, 2003). An artificial group identity is created that is not at all related to the substantive focus of the choices made in the later experiment. Although subjects were able to communicate during the stage of group formation, these papers assume that the group identity is independent of uncontrollable social interaction effects that could influence the behavior of individuals within and outside the group.

In our experimental design, we change the artificial quality of group identity and link group identity more closely to the tasks individuals have to solve in the experiment. Moreover, we indirectly motivate individuals to draw conclusions about common values and social preferences shared in their group. This stands - for example - in contrast to the philosophy of group formation as used in Chen and Li (2009), where an individual group member cannot draw conclusions about the pro-social preferences of his or her group members (and the members of the other group) following the process of group formation and later in the experiment. We use the attitudes expressed by individuals on environmental and recycling issues for grouping them in order to test whether commonly shared values influence individual behavior. As recycling activities lead to reduced human impact on the environment, similar environmental attitudes should become a common ground of every club. The process of group formation and group identity reinforcement which subjects went through in the experiment can be summarized as follows:

- At the beginning of the experiment, participants were asked to complete a questionnaire consisting of ten questions (see Appendix). Items were obtained from Fernández-Manzanal et al. (2007) and adapted to our purpose. For every question, participants expressed their attitude toward the environment on a 6-point Likert scale. It is important to note that the subjects were unaware that we would use this information for later grouping.

- After playing treatment 1, the group formation took place. The answers to the ten different questions were aggregated to an average index in order to rank all *N* society members accordingly and split societies into clubs of *N*/2 members with "club 1" in which all members appreciate the environment at least as much as any member of the other "club 2."
- Subjects received three pieces of information: First, they were informed of being part of a club with all members having similar attitudes toward the environment. The denomination of both clubs ("club 1" or "club 2") was chosen to be neutral. Second, they were informed about the index of their own environmental attitudes as stated at the beginning of the experiment. And third, they were also notified about the average index of the members of the society they were assigned to. Consequently, the group membership was not explicitly mentioned, but could be derived from the given information.
- In order to stimulate the feeling of belonging together, the clubs competed in a slider-game. The task of the club members was to realize as many as possible proper slider settings within a minute. In each society, two clubs competed against each other. The members of the winning club received a potential extra payment, paid out at the end of the experiment. Information on whether the own club won was revealed directly after the slider-game.

6.4.4 Incentive mechanism

In both treatments and after every round, the individual contributions to the recycling system of all players of the group, as well as their payoff, were communicated on the screen (see Appendix). The experiment was incentivized by a random lottery mechanism. Subjects were told that at the end of the experiment all six members of one randomly selected society would receive the payoff they achieved in a randomly selected round, with one token equaling 1 EUR. In total, 178.80 EUR were paid out: 96.00 EUR in treatment 1; 76.80 EUR in treatment 2, and 6 EUR in the process of identity reinforcement. On average, every participant received 4.97 EUR for participating in the experiment.

6.4.5 Hypotheses

The experimental design described above focuses on two research questions: The first question relates to potential effects of introducing a multiple institutional designs in the case of recycling as opposed to a pure public good structure. The social return to both types of investment options

is the same. However, such kind of financial equivalence disregards expectation formations about the behavior of other subjects. Assuming that there is uncertainty about the behavior of all other players, investing in the club good might become more attractive due to the larger MPCR. The second question is directed to possible in-group effects when clubs discriminate between certain attitudes. More specifically, we investigate whether there are possible group effects if the composition of the groups is based on similar environmental attitudes of their members as suggested by Brekke et al. (2010), who identified a certain peer group effect. We explicitly assume that participants with stronger environmental attitudes have stronger pro-social preferences and therefore contribute higher shares of their endowments. Hence, our hypotheses can be expressed as follows:

- **Hypothesis 1** The sum of individual contributions to recycling systems increases when several alternatives are offered.
- **Hypothesis 2a** The contributions to the club good are higher in the club whose members share stronger attitudes toward the environment.
- **Hypothesis 2b** The total contributions (i.e., the sum of contributions to the club good and the public good) are higher in the club whose members share stronger attitudes toward the environment.

6.5 Results

The results section is structured as follows: After first presenting the more general descriptive results and second analyzing the composition of the clubs, we turn to the implications of the previously formulated hypotheses. In contrast to the instructions, we follow the economic language and henceforth speak of public goods (PG) and club goods (CG).

6.5.1 Descriptive results

Table 6.2 provides an overview of the average contributions to the PG, the CG, and total contributions as well as average payoffs. Results are separated by treatment and round: The values for round 1 represent average contributions and payoffs of all participants in the first round only, while the values for round 1–10 refer to average contributions of all participants in all rounds. Results show that in treatment 1, participants contributed 38% of their endowment to the recycling system, generating an average payoff of 13.04 EUR per round. In treatment 2, total contributions (44%) and payoffs (13.52 EUR) are higher as almost 68% of contributions

are made using the CG. Except for the public good in treatment 2, average contributions in the first round exceed the mean of all ten rounds and indicate a decreasing trend over time.

Treatment	**Treatment 1**		**Treatment 2**	
Round	**1**	**1–10**	**1**	**1–10**
PG contributions	4.25 (3.57)	3.80 (3.00)	1.22 (1.90)	1.41 (1.94)
CG contributions	–	–	3.89 (3.59)	2.99 (2.63)
Total contributions	4.25 (3.57)	3.80 (3.00)	5.11 (3.62)	4.40 (3.28)
Payoffs	13.40 (3.55)	13.04 (2.79)	14.09 (3.41)	13.52 (2.83)

Table 6.2: Average contributions and payoffs and standard deviations.

Notes. Standard deviations in parentheses.

Figure 6.1 allows for a closer examination of total contributions in both treatments. It has to be kept in mind that using the within-subjects design, all players play treatment 2 after having played treatment 1. In round one, total contributions are slightly lower in treatment 1 than in treatment 2 (4.25 versus 5.11 units). A tendency of decreasing contributions with a minimum of about 2.8 units in round ten can be observed for both treatments. In treatment 2, the average total contributions of participants to recycling systems over all rounds are about 16% higher (4.40 units) than in treatment 1 (3.80 units). Yet, the curve progressions are rather similar and declining—a standard result in public good experiments (Ledyard, 1995).

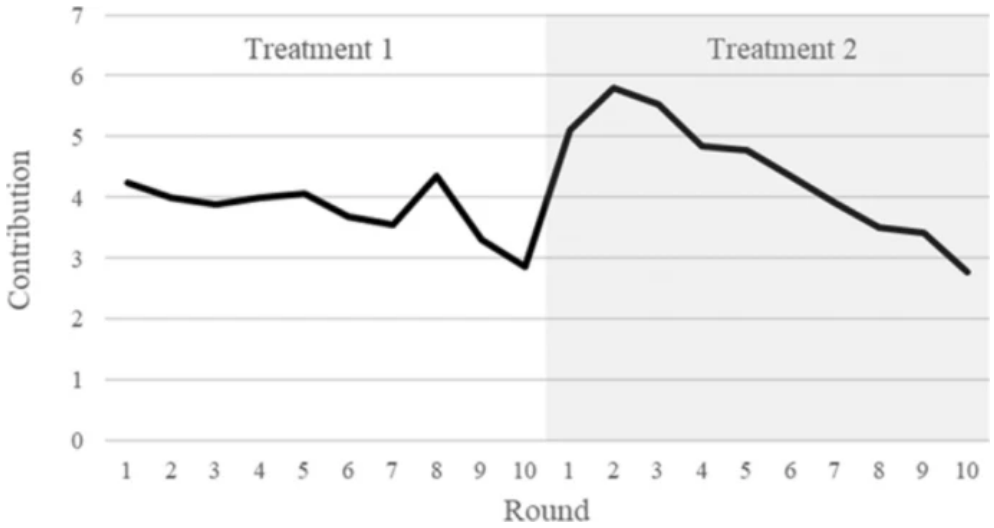


Figure 6.1: Average total contributions in both consecutive treatments.

6.5.2 Composition of the clubs

As described above, the allocation to the different clubs is based on individuals' environmental attitude. This was determined by using an average from answers to ten questions (see Appendix)

using a 6-point Likert scale. The answers ranged from 0 ("Environment is not important at all to me") to 5 ("Environment is very important to me.") The average environmental attitude of members in Club High was 3.67, and 2.08 in Club Low, reflecting a significant difference. Further statistical analyses show no other significant differences regarding the composition of the groups. Using a Chi-Square test, we checked for differences in terms of subjects' gender, income, secondary employment, nationality, and religious affiliation. In the slider-game, which was conducted between the two treatments, members of Club High solved an average of 3.77 out of 10 tasks right, while Club Low received a similar result of 3.22 correct answers. Furthermore, the age of the undergraduate students does not differ considerably.

6.5.3 Testing hypothesis 1

Figure 6.1 depicts the overall contributions of both treatments over time. Bearing in mind that the order and financial attractiveness of contributions are similar in both treatments, this graph reveals interesting results. Instead of continuing a declining trend after treatment 1, the overall contributions surge to a new maximum level in the first round of treatment 2. The peak is reached in round 2 where 5.8 out of 10 units are contributed to both goods. After that, contributions decrease again.

For a statistical analysis of differences in contributions between the two treatments, we use the Wilcoxon Sign-Rank test and define the threshold of statistical significance as $p = 0.05$. This nonparametric test is used due to the absence of normally distributed values (e.g., due to differences between mean and median values as shown in Table 6.3) and since it is applicable to small sample sizes. The test compares participants' contributions in the respective rounds of both treatments to check the equality of the central tendencies. Table 6.3 provides the results of this test for all participants as well as separated by the clubs. Although participants were not aware of the separation into two clubs during the first treatment, their contributions can be traced back and evaluated ex-post.

	Club High	Club Low	All
Total contributions in Treatment 1	4.08 (3.05)	3.52 (3.15)	3.80 (3.10)
Total contributions in Treatment 2	5.51 (5.00)	3.29 (2.95)	4.40 (4.25)
Wilcoxon Sign-Rank	$p = 0.006$	$p = 0.421$	$p = 0.112$

Table 6.3: Comparison of contributions across treatments and clubs.

Notes. Median in parentheses.

Results show that hypothesis 1, which predicts an increasing sum of individual contributions to a recycling system after offering several alternatives, cannot be proven true for all rounds and participants ($p = 0.112$). A closer examination shows significant differences for the first three rounds only (round 1: $p = 0.048$, round 2: $p = 0.002$, and round 3: $p = 0.007$).

Nonetheless, we find meaningful results when differentiating between the different clubs. Contributions of Club High members significantly increase in treatment 2 compared with treatment 1 ($p = 0.006$), while no significant differences can be shown for Club Low. At this point, we refer to the relevance of the composition of the two clubs which will be discussed in more detail later.

6.5.4 Testing hypotheses 2

The second hypothesis asks about possible differences in the subjects' behavior in relation to their association with the clubs in treatment 2. Unlike in treatment 1, the subjects now know that they have been assigned to a club according to their environmental attitude and have their additional CG recycling option. As this comparison is based on a between-subjects level, the hypothesis is tested by using the nonparametric Mann–Whitney U criterion. Again, this test tends to be a good option in the absence of normally distributed values and is applicable to our sample sizes. We define the threshold of statistical significance as $p = 0.05$.

Table 6.4 allows for a comparison of average contributions and payoffs in both clubs in treatment 2 in order to test hypotheses 2. Results show that Club High members made significantly higher contributions to the CG than Club Low members in treatment 2 ($p = 0.012$). Hence, these results support hypothesis 2a. Regarding hypothesis 2b, stating that total contributions are higher in Club High, the results are not so clear: Despite differences in the total contributions of the two clubs, the Mann-Whitney U test results in $p = 0.059$ and does - though marginally - not attest a significant increase. Overall, Club High members realized significantly higher payoffs as compared with Club Low members in treatment 2 ($p = 0.037$).

Besides, no significant differences in the two clubs' contributions to the PG and, more interestingly, to the PG in treatment 1 can be detected (see gray bottom of Table 6.4).

	Club High	**Club Low**	**Mann–Whitney U**
Treatment 2			
PG contributions	1.39 (0.35)	1.44 (0.30)	$p = 0.938$
CG contributions	4.12 (3.95)	1.86 (1.30)	$p = 0.012$
Total contributions	5.51 (5.00)	3.29 (2.95)	$p = 0.059$
Payoffs	14.45 (13.27)	12.59 (12.09)	$p = 0.037$
Treatment 1			
Total contributions	4.08 (3.05)	3.52 (3.15)	$p = 0.521$
Payoffs	12.75 (12.73)	13.32 (13.15)	$p = 0.628$

Table 6.4: Comparison of contributions and payoffs in club high and club low.

Notes. Median values in parentheses.

A further impression of the composition of the contributions over the rounds can be obtained by consulting Figs. 6.2 and 6.3. They show that the contributions to the PG are rather stable and do not differ significantly for the two clubs (Mann-Whitney U: $p = 0.938$). Differences in the total contributions are thus mainly caused by contributions to the CG. Furthermore, the decreasing contributions in the second treatment - in contrast to the first - are mainly due to contributions to the CG. Reasons for this different contribution behavior can therefore be linked to the different club compositions.

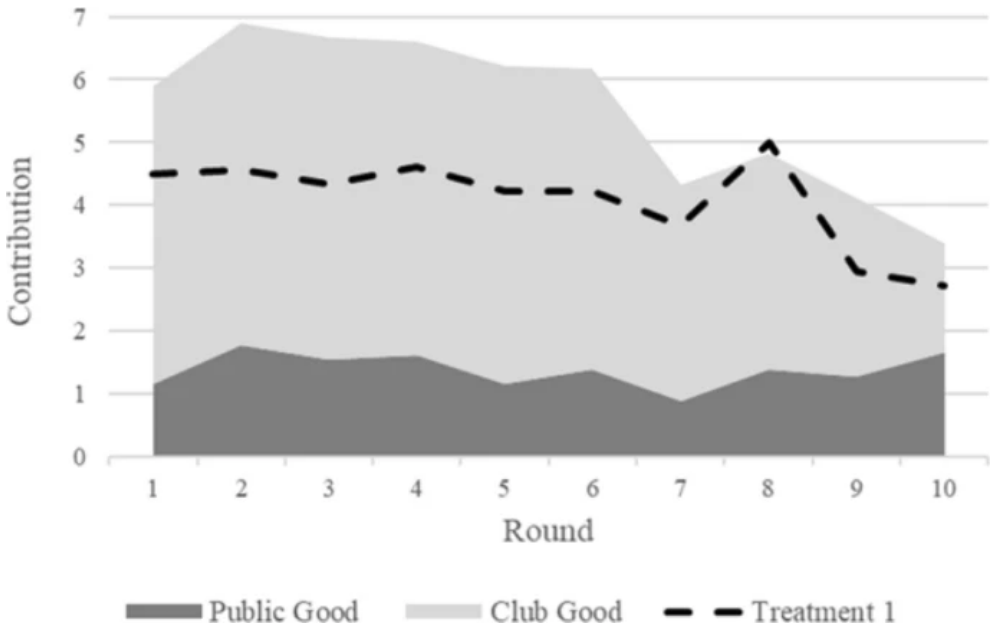


Figure 6.2: Contributions to public good and club good in treatment 2 for club high.

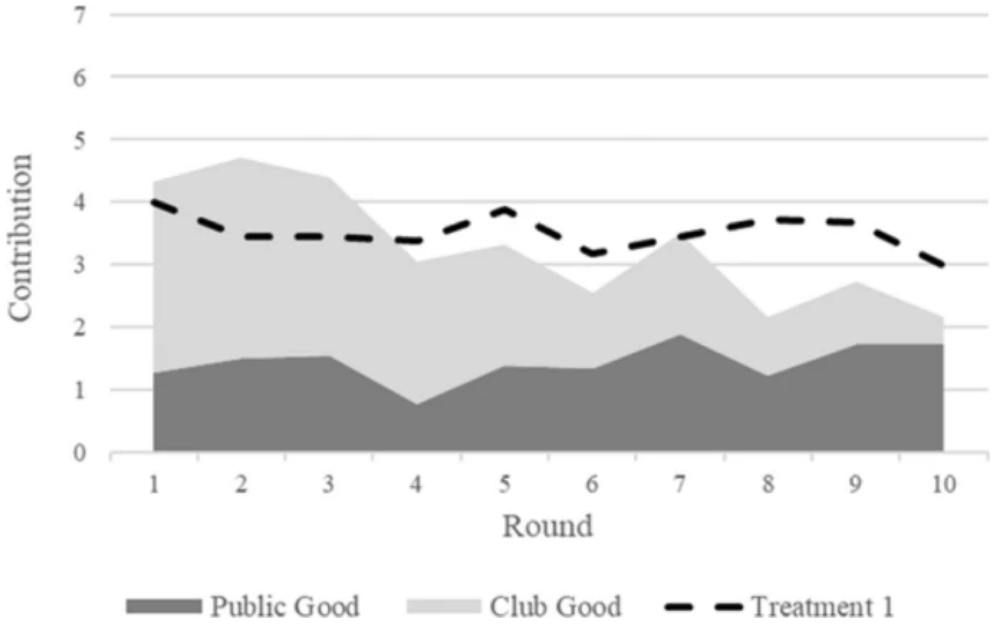


Figure 6.3: Contributions to public good and club good in treatment 2 for club low.

6.6 Conclusion

The results of the experiment reported in this paper primarily point to the fact that subjects show a relatively stable degree of cooperation in both consecutive treatments. On average (over all rounds), subjects invested 38% of their endowment in treatment 1 and 44% in treatment 2. At the same time, a standard result of experimental economics which postulates a clear reduction in cooperation from round to round in repeated games (Ledyard, 1995) is confirmed in this study. Since both treatments were played in succession (with the same set of co-players in fixed societies) without varying their order, it is of course not unlikely that the round-related decline in cooperativeness also took place across both treatments. However, in an experiment by Andreoni (1988), a public goods game (PGG) is played over ten rounds and then restarted with the same group composition. The restarted PGG is repeated only three times which is, however, not known by the subjects. The results show that the contributions in these three rounds are only minimally lower (a 0.2 decline of contributions in round 1 of the restarted PGG) than in the first PGG. In our experiment, there is also no reduction in cooperation when comparing both treatments. There are no statistically significant differences in total contributions to the recycling systems. Against this background, it is surprising that the average total contributions in treatment 2 exceed contributions in the first treatment by 16%. When we consider only the first round of both treatments, the contributions in treatment 2 are 20% higher compared with treatment 1. This indicates that by introducing the club good option an activation of cooperation could take place, which potentially counteracts the trend of reducing cooperation in repeated games.

The composition of the clubs in the second treatment, on the other hand, seems to induce a significant effect on individuals' contributions to the recycling system. On average, the members of the club with stronger environmental attitudes (Club High) invested 41% of their endowment in the club good, while members of the other club (Club Low) only invested about 19%. In Chen and Li (2009), an individual group member cannot conclude the pro-social preferences of his group members (and the members of the other group) through the process of group formation. However, participants in our experiment may assume that stronger environmental attitudes are correlated with stronger pro-social preferences. If a participant is aware of belonging to the group with the lower environmental attitudes, he might assume that the average contributions in his group are lower than that in the group with the stronger environmental attitudes. This would, of course, have an impact on whether he decides to invest in the club good. Although contributions to the club good decrease from round to round, which is common in repeated games, this observation is in line with the result of Brekke et al. (2010). They state that households who firmly believe that their recycling attitudes are shared within their group derive a greater sense of responsibility for participating. An explanation of this experimental group effect could be that it stimulates a sense of belonging, which allows subjects to derive conclusions about their own responsibility. Quite remarkably, contributions to the public good do not differ much between these clubs and do not decrease from round to round.

A complementary explanatory approach can be made with the concept of identity economics from Akerlof and Kranton (2000). In this context, the two clubs represent social categories and their social norms serve as a guideline for the disposal decision. All participants are aware that they have been divided into groups of high and low environmental attitudes and can draw conclusions about which group they are in. Positive recycling behavior tends to be the norm for the group with positive environmental attitudes. This forecast is also reflected in our results: Club Highs' contributions to their CG is higher. At the same time, in the existing dilemma situation, this results in a separation from the other club which is subjected to different norms and thus reflects a different recycling behavior. For policymakers, this implies that relevant standards should be strengthened in all social subgroups.

As described, a central observation of the experiment refers to the total contributions that individuals make to the offered recycling systems. Note that the introduction of the club good option in the second treatment is an extension of the options for action without changing the material incentive structure of the first treatment. The introduction of the additional club good

option is therefore merely a change in the social context of the individual decision. However, this variation in the social context led to an increase in the total contributions on average. While the individuals assigned to the respective "Club Low" do not show a statistically significant change in their total contributions, the total contributions of the individuals assigned to the respective "Club High" increase significantly. Therefore, a reasonable policy implication is to localize recycling systems and make their benefits local public goods.

It is difficult to answer whether the results observed in the experiment are due to the creation of exclusion possibilities in the case of club goods, or whether this is rather due to the stronger environmental attitudes of the individuals in Club High. In fact, the influence of these two factors cannot be separated. Environmentally friendly individuals did not behave significantly different from the less environmentally friendly individuals in the first treatment. Statistically significant differences in the behavior of these two groups arise solely in the second treatment with the introduction of the club good option. It is only the change in the social context of individual interactions that leads to the fact that the individuals with stronger environmental attitudes make higher contributions than the individuals with weaker environmental attitudes. In order to discriminate between the two factors "introduction of a club good option and excludability" and "environmental attitudes," we will vary the experimental design in further experiments and test for the influence of alternative grouping variables when introducing the clubs. In addition, future investigations should rather be conducted as between-subjects designs in order to exclude order effects of the treatments. If the groups' influence on the cooperative recycling behavior in future investigations can be substantiated, an interesting extension of these experiments is to include endogenous group processes, for example through elements of gamification.

With regard to the further procedure, the question may be raised whether further insights into individual recycling behavior could also be obtained from simulations instead of experiments. Basically, experiments and simulations represent complementary methods to analyze the connection between individual behavior and social results (Dawid and Delli Gatti, 2018). However, the use of (agent-based) simulations is based on concrete assumptions about the structure of individual preferences and the decision-making behavior of individuals. Here, deviations from the standard assumptions of microeconomics can be modeled and simulated using stochastic processes as well. The essence of experiments, on the other hand, is that the individual behavior is empirically observed in view of a given incentive structure and is not

given exogenously. The question of whether individuals feel motivated to change their behavior in repeated interactions when the social context of interaction is varied cannot be answered with simulations. However, simulation programs can be used to transfer knowledge gained from experiments to other model relationships.

The experiment presented in this paper was initially conceived as a pilot study which aimed to obtain information for the design of a broader study. The results, especially those regarding hypothesis 1, are to be interpreted with caution, since the design of the experiment did not allow for a variation of the treatment order: This uncontrollable sequence effect is likely to have been at the expense of cooperation in treatment 2. However, the results of the introduction of the club good option turned out to be even more interesting, as this seems to induce a significant effect on cooperation behavior.

6.7 Appendix

6.7.1 Questions used for club formation (based on Fernández-Manzanal et al., 2007)

1. I am willing to consume less and to forgo some comforts if it helps to protect the environment.
2. I am willing to spend more money in order to purchase a recyclable packaging.
3. I like spending time in nature to understand the environment in which I live.
4. I am trying to find out how my behavior influences the environment.
5. If I have to choose between a new highway and a conversation area, I will choose the conversation area.
6. If the public transport system was more efficient, I would prefer it to the car.
7. We should try to conserve plants and animals, even though it is expensive.
8. Environmental education activities for children are important.
9. Our laws should be changed in order to reduce pollution of the environment.
10. Universities should schedule more environmental issues.

6.7.2 Screenshot

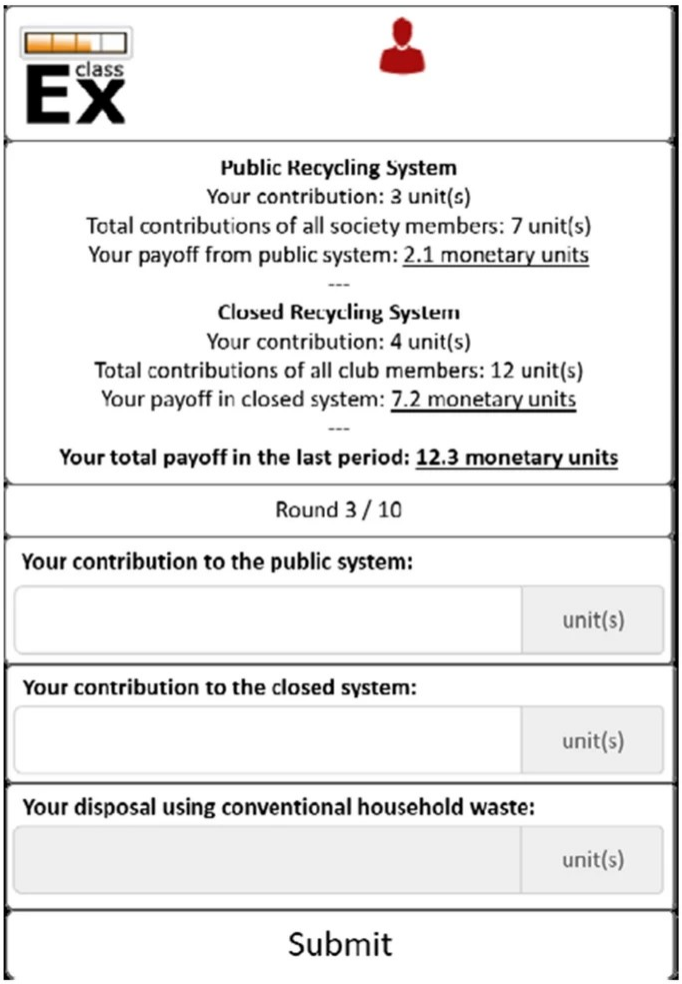


Figure 6.4: Decision screen in classEx (example).

7. Combating climate change: Is the option to exploit a public good a barrier for reaching critical thresholds? Experimental evidence

This chapter has appeared as:

Cloos, J. & Greiff, M. (2021). Combating Climate Change: Is the Option to Exploit a Public Good a Barrier for Reaching Critical Thresholds? Experimental Evidence.

Abstract: The achievement of collective climate targets is hampered by a large number of factors. Most obvious is the conflict between self-interest and group interest at both the intra- and intergenerational level. Several experimental studies examine the effects of factors such as wealth heterogeneity, varying thresholds, or time discounting on the probability of achieving a collective climate target. In these experiments, participants act as a group and can invest money in a collective group account over a fixed number of rounds. If the group account is below a threshold after the last round, the members of a group usually lose a large proportion of their potential assets. However, in the real world, agents can not only invest in public goods, but also exploit them. We therefore study cooperation dynamics in a threshold climate change experiment in which group members can not only contribute money into their group account, but also take money out of it. We induce endowment heterogeneity by simulating the contribution decisions in the first rounds of the experiment and vary the loss rate between treatments. Our results show no significant differences between give and give-take treatments. Consistent with the results of previous studies, we find that with a lower loss rate, less groups reach the threshold. (JEL: C92, D74, D81, H41, Q54)

Keywords: climate change, experiment, public goods game, threshold public goods game, exploitation

7.1 Introduction

Already by 2017, human activities had resulted in the warming of the global climate by about 1°C compared to pre-industrial levels. If the current rate of global warming is maintained, a global warming of 1.5°C will be reached between the years 2030 and 2052. Negative consequences of climate change are already evident in different regions and will increase significantly if global warming rises to 1.5 or even 2°C (IPCC, 2018). While investments aimed at effectively mitigating global warming are undertaken by individuals or groups of individuals, the rewards from climate-protecting investments benefit everyone. Consequently, there is a free-rider problem associated with climate protection. Individual actors have an incentive not to contribute to climate protection and to benefit from the efforts of others. At the same time, the collective benefit would be largest if investments are high enough so that catastrophic events can be avoided (Gollier and Tirole, 2017). Investments in climate protection are further characterized by the fact that the resulting benefits will largely be realized in the future. One of the key challenges facing humankind is therefore how to achieve intra- and intergenerational cooperation to combat climate change (Nordhaus, 2019).

The underlying decision situation can be described as a public good game (PGG) in which the entire world population forms the set of players (e.g., Milinski et al., 2008). Alternatively, one could argue that this decision situation resembles a group of countries engaged in climate negotiations (e.g., Tavoni et al., 2011). Individual contributions to the public good correspond to actions that reduce or limit CO_2 emissions.

In contrast to linear PGGs (see e.g., Zelmer, 2003; Ledyard, 1995), the social optimum in this global PGG does not consist in all actors investing the entire amount available to the public good. Rather, the problem of global warming can be modeled as a threshold PGG in which the public good is provided only if the sum of contributions exceeds a given threshold. For example, the goal of the Paris agreement is to keep the increase in global average temperature this century below 2°C compared to pre-industrial levels (United Nations, 2015). Based on this threshold, one could determine the aggregate contributions (i.e., worldwide reduction in CO_2 emissions) needed to reach the target.

The collective-risk social dilemma (CRSD, Milinski et al., 2008) is one variant of a threshold PGG. In the CRSD, a fixed number of participants interact over a finite number of rounds. Within each round, participants simultaneously choose their contributions. If the sum of contributions at the end of the last round exceeds a predefined threshold, every participant keeps

her non-contributed wealth. Otherwise, each participant loses a fraction of her remaining wealth with a given probability.

In recent years, Milinski et al. (2011, 2008) and a number of related studies (Waichman et al., 2018; Brown and Kroll, 2017; Dannenberg et al., 2015; Jacquet et al., 2013; Tavoni et al., 2011) have used the CRSD or extensions of it to experimentally investigate under which conditions humans can succeed in preventing dangerous climate change. We will review these studies in section 7.2.

Countries[41] can contribute to climate protection by, for example, reducing their CO_2 emissions through investments in renewable energies or by enforcing stricter regulations. Climate-protecting actions have the characteristic that most of the benefits arising from them only become evident in the medium to long term (intergenerational). However, countries can also intensify the existing climate problem by, for example, clear-cutting forest areas to create space for new farmland (Mitchard, 2018; Rochedo, 2018) or exploiting new coal mining areas (Jakob et al., 2020; Blondeel and Van de Graaf, 2018). These kinds of climate worsening activities have the characteristic that they are implemented in order to boost short-term economic growth, independent of whether the threshold will be reached or not.

The remaining question is in how far climate-damaging activities of individual actors motivate former climate-friendly actors to reduce their climate-friendly efforts or even carry out climate-damaging activities themselves. A further related question is how former climate-friendly actors behave when it becomes apparent that threshold values can no longer be reached at a certain point in time.

However, in all existing studies on the CRSD, participants only have the option of investing money to achieve the climate target. Yet, the above-mentioned examples on climate damaging activities illustrate that actors can not only free-ride but also actively exploit other actors' contributions. It is unclear whether the results from existing experiments also apply to situations in which exploitation is possible. Moreover, by neglecting the possibility to exploit, existing studies might lead to overly optimistic results.

Our experiment addresses this gap by examining the question how the dynamics of achieving the climate target change when there is a take option, in addition to the give option. The main

[41] Although we use the term countries in this study, it is equally applicable to smaller entities such as counties, organizations, or individuals.

research question of this paper is therefore whether the possibility of exploiting a public good and the observation of such behavior is a barrier to effective cooperation in the CRSD.

Our experiment consists of 20 rounds, though participants only make active contribution decisions in rounds 11 to 20. The experiment contains six treatments which differ according to participants' action sets, the potential loss rate, and participants' initial wealth distribution. In three Give-treatments (G-treatments), participants can only contribute non-negative amounts to the group account. In three Give-Take-treatments (GT-treatments), participants can either contribute non-negative amounts to the group account or withdraw amounts from it. We induce wealth heterogeneity by simulating the contribution decisions in the first 10 rounds of the experiment.

Our results show that the introduction of a take option makes it more difficult for groups to reach the threshold. In G-treatments, more groups reach the threshold than in the corresponding GT-treatments, although the effect is not significant. Consistent with previous literature, we find that a higher potential loss rate results in more groups reaching the threshold. We observe that groups with homogeneous wealth reach the threshold more easily than groups with heterogeneous wealth, but again, the difference is not significant. Taken together, our results show that extending the range of the action set to the negative domain and inducing heterogeneous wealth complicates coordination within a group.

The remainder of this paper is structured as follows: In section 7.2 we discuss the related literature. We describe the experimental design, procedures and hypotheses in section 7.3. In section 7.4, we present the results of our experiment. We discuss the results, point to a number of limitations and conclude in section 7.5.

7.2 Related literature

7.2.1 The collective-risk social dilemma (CRSD)

As mentioned above, the CRSD was introduced by Milinski et al. (2008). In their experiment, groups of six participants play the CRSD for ten rounds. At the beginning of the game, each participant has an endowment of € 40. Within each round, participants simultaneously decide whether to contribute € 0, 2, or 4 into a group account.

All group members know that at the end of the ten rounds they will receive their non-contributed endowment only if the sum in the group account reaches or exceeds the threshold of € 120. On average, this threshold corresponds to a contribution of € 2 per participant and round. If the

threshold is not reached after ten rounds, the non-contributed endowment will be lost with a probability of p=90% (T90), p=50% (T50), or p=10% (T10), depending on the treatment. Participants will not receive any payment in case of loss.

In case each participant contributes her fair share of € 2 per round, the threshold is exactly reached and each participant receives a payoff of € 20. This holds true for all treatments. However, if all participants free ride (i.e., invest nothing), the threshold is not reached. In this case, the expected payoff for each participant is € 40 * (1-p).

As the expected payoff from free riding decreases in p, one would expect that the threshold is reached more often when p is high. This is exactly what the results show. In T90, the threshold was reached by 5 of the 10 groups (mean € 118.2). Only one group reached the threshold in T50 (mean € 92.2) and no group reached the threshold in T10 (mean € 73.0). Although almost all participants in each treatment contributed € 2 in round 1, the willingness to contribute decreased in subsequent rounds, especially in T50 and T10.

7.2.2 Extensions of the CRSD

The CRSD has spurred a large and still-growing literature in which several extensions of the CRSD have been investigated. In this section, we discuss the extensions which are most relevant for our experiment, namely CRSDs in which participants have an operating fund and an endowment, and CRSDs with wealth heterogeneity. Other extensions of the CRSD investigate the effects of communication (Tavoni et al., 2011), uncertainty about the threshold (Dannenberg et al., 2015; Barrett and Dannenberg, 2014, 2012), heterogeneity in the expected loss (Brown and Kroll, 2017; Burton-Chellew et al., 2013), and heterogeneity in the wealth distribution and loss probabilities (Waichman et al., 2018).

Milinski et al. (2011) is the first study in which participants have an operating fund and an endowment. The authors examine the effect of intermediate climate targets and wealth heterogeneity on group cooperation. As in Milinski et al. (2008), participants interact in groups of six for ten rounds and the threshold is € 120. Contributions to the group account are paid out of the operating fund, which should mimic participants' wealth which they can use to cover their living expenses. In contrast to Milinski et al. (2008), participants receive their remaining amounts from the operating fund after the end of the experiment, even if the threshold is not reached. The endowment does not change during the game but will be lost with a probability of 90% at the end of the game if contributions fall short of the threshold. The endowment can be seen as assets that are negatively affected by medium- and long-term climate changes.

Within this setting, Milinski et al. (2011) study the effect of wealth heterogeneity by distinguishing between “rich” and “poor” participants. "Rich" participants have an operating fund of € 40 and an endowment of € 60, while "poor" participants have € 20 and € 30, respectively. Treatments with exclusively "rich" participants, with exclusively "poor" participants, and with a mixed number of three "rich" and three "poor" participants are tested. The results show that all "rich" groups, no "poor" group, and 60% of the mixed groups managed to reach the threshold. Interestingly, "rich" and "poor" participants in the mixed groups showed no different contribution behavior than in the groups with only "rich" or only "poor" participants.

Using a similar experimental setting, Jacquet et al. (2013) investigate the effect of time-dependent discounting on contribution behavior. In their experiment, the non-invested operating fund is paid out to the participants directly after the experiment. In case the threshold is reached, the endowment of 45 € is paid out after one day (treatment 1), after seven days (treatment 2), or it is invested in planting oak trees (treatment 3). If a group falls short of the threshold, the group members’ endowments are destroyed with a probability of 90%. While in treatment 1, 7 out of 10 groups reached the threshold, only 4 out of 11 groups managed to reach the threshold in treatment 2, and none of the 11 groups in treatment 3.

In a review article, Dannenberg and Tavoni (2016) compare the results of experimental climate change games with the results of related studies from evolutionary game theory. Based on the findings of the reviewed studies, Dannenberg and Tavoni (2016, p. 95) come to the general conclusion “…that the expected loss of crossing the threshold and uncertainty about the threshold are the most important determinants of collective action. The threshold’s role as a catalyzer of cooperation is hindered when uncertainty and especially ambiguity about its location is introduced. Wealth inequality and the credibility of the pledges constitute further difficulties.”

The results from Milinski et al. (2011) and from Brown and Kroll (2017) provide no clear evidence for a negative effect of wealth heterogeneity. However, the findings of Tavoni et al. (2011) and Burton-Chellew et al. (2013) suggest that with wealth heterogeneity, groups are less likely to reach the threshold. In contrast to these results, Waichman et al. (2018) find that heterogeneities can facilitate coordination. Although the main contribution of our study is the extension of participants' action sets by allowing negative contributions, we also conduct treatments with homogenous and heterogeneous operating funds.

7.2.3 Related literature on give and take-options

Our experimental design is also related to experimental studies that investigate the effect of extending participants' action sets in PGGs and dictator games. Starting with Andreoni (1995), a number of experimental studies (Gächter et al., 2017; Dufwenberg et al., 2011; Park, 2000; Sonnemans et al., 1998) have shown that cooperativeness is higher in positively framed PGGs than in negatively framed public bad games. For a repeated PGG, Khadjavi and Lange (2015) show that average contributions in a treatment that allows both positive (give) and negative (take) contributions are not significantly different from a strategically equivalent pure give treatment. However, the effect of a simultaneous give-take option has not been studied yet in a CRSD.

Krupka and Weber (2013), Bardsley (2008), and List (2007) conduct dictator games in which the dictator can either give or take away money from the recipient. In Krupka and Weber's (2013) standard dictator game, dictators are endowed with US$ 10 and recipients with US$ 0, and dictators can give between US$ 0 and US$ 10. In a "bully" variant, both dictators and recipients are endowed with US$ 5 each, and dictators can choose between giving up to US$ 5 or taking up to US$ 5. Note that both decision situations are strategically equivalent because they have the same number of actions which result in the same set of outcomes. Two results are striking: First, the equal split, in which both the dictator and the recipient receive US$ 5 each, occurs more often in the "bully" variant. Second, allocations in which the recipient receives less than US$ 5 are more frequent in the standard dictator game. Moreover, Krupka and Weber (2013) conduct an additional experiment in which they elicit norms about the social appropriateness of the different actions from the standard and "bully" dictator games. The appropriateness ratings are in line with the results. That is, in the "bully" variant, the equal split is perceived as more appropriate than in the standard dictator game, and taking money is perceived as less appropriate than giving money, even though both actions lead to identical payoffs.

7.3 Experimental design and hypotheses

7.3.1 Experimental design

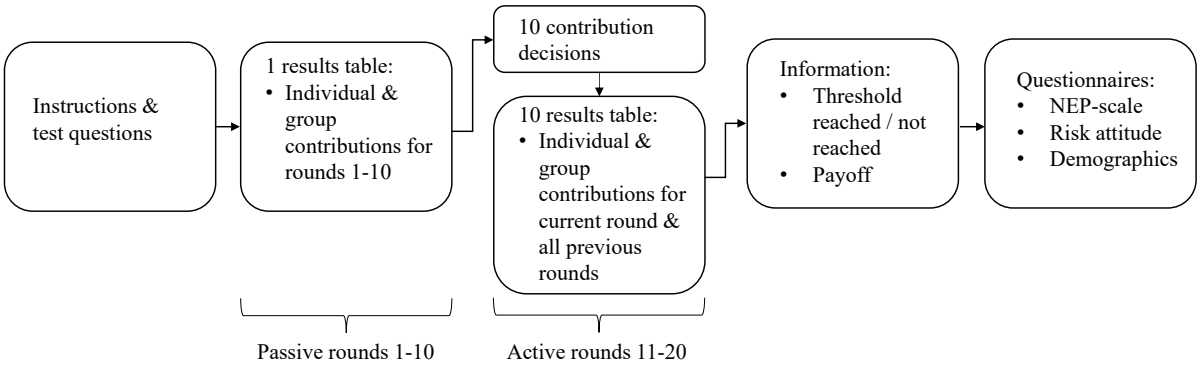


Figure 7.1: Sequence of the experiment.

Our experimental design is shown in Figure 7.1. Our experiment is a CRSD with N=6 participants per group and *T*=20 rounds. After the instructions (Appendix A.1, section 7.6.1) and test questions, there are ten passive rounds followed by ten active rounds of the CRSD (described in detail below). Starting at the end of round 10, participants are informed about all group members' contributions within each round, each group members' aggregate contribution, and the value of the group account (see the results table in the experimental instructions in Appendix A.1, section 7.6.1). In addition, all participants receive information on the remaining operating fund of each group member up to the current round, and the average contributions of each group member up to the current round. To ensure anonymity, all participants are given a pseudonym. The pseudonyms correspond to the names of moons and dwarf planets from the solar system and are retained by the participants throughout the entire experiment. At the end of round 20, participants are told whether the threshold is reached or not and receive information on their final payoff. The experiment ends with a post-experimental questionnaire on environmental and risk attitudes and demographic characteristics. In the following, we describe each part of the experiment in more detail.

Contribution decisions in rounds 1-20: Within each round $t = (1, \dots, T)$, all participants simultaneously chose their contributions c. Let $c_{i,t} \in A_i$ denote the contribution of participant i in round t. At the beginning of round 1, the group account is empty, i.e. $C_0 = 0$. Let $C_t = \sum_{i=1}^{n} \sum_{\tau=1}^{t} c_{i,\tau}$ denote the value of the group account at the end of round t. Let $OF_{i,t}$ denote participant i's operating fund at the beginning of round t. Since contributions are paid out of the operating fund, the operating fund evolves according to $OF_{i,t+1} = OF_{i,t} - c_{i,t}$. The initial value of the operating fund is given by $OF_{i,1} = 160$.

We denote participant *i*'s endowment by $e = 120$. The initial endowment is the same for all participants and does not change during the game, hence we can omit all subscripts. If, at the end of the game, the group account equals or exceeds the threshold $S = 480$, the endowment is unaffected. Otherwise, a fraction p of the endowment is destroyed. Note that p is not a probability but a fixed loss rate.[42] Since initial endowments are identical, the expected value of the endowment at the end of the game is identical for all participants.

At the end of the game, each participant's payoff consists of her operating fund (OF_i) minus her contributions over all 20 rounds plus an additional payoff (i.e., the endowment), which depends on the value of the group account. Equation (1) summarizes the payoffs.

$$\pi_i = \begin{cases} OF_{i,T+1} + (1-p)e & \text{if } C_T < S \\ OF_{i,T+1} + e & \text{if } C_T \geq S \end{cases} \quad (1)$$

Participants' action sets are treatment-dependent and allow either only non-negative (give) contributions $A_i^G = \{0, 1, 2, 3, 4, 5, 6, 7, 8\}$, or giving and taking, $A_i^{GT} = \{-3, -2, -1, 0, 1, 2, 3, 4, 5, 6, 7, 8\}$. Within each treatment, action sets are identical for all participants. In the Give-Take-treatments (GT-treatments), participants can take money out of the group account, possibly reducing the likelihood that the threshold will be reached. This option should reflect the real world climate-damaging possibility of exploiting natural resources. By allowing participants to take money out of the group account, they can increase their operating fund. This is not possible in the Give-treatments (G-treatments). Note that in all treatments, at the end of the game, participants receive the value of the operating fund $OF_{i,T+1}$ regardless of whether the threshold is reached or not.

Heterogeneity: Our experiment includes treatments with homogeneous operating funds (HOM-treatments) and treatments with heterogeneous operating funds (HET-treatments). Similar to Tavoni et al. (2011), we implement homogeneity and heterogeneity by simulating participants' decisions in rounds 1 to 10, that is, contributions in rounds 1 to 10 are determined by the computer.

In the HOM-treatments, each participant contributes a total of 40 over rounds 1 to 10 (passive rounds). Hence, each participant starts round 11 with an operating fund of $OF_{i,11} = 120$.

[42] In most variants of the CRSD p is a probability. The parameter p is a fixed loss rate in treatments T3 and T6 in Brown and Kroll (2017) and in treatment Certainty in Dannenberg et al. (2015).

In the HET-treatments, within each group, three participants each contribute a total of 20 over rounds 1 to 10, and the three remaining participants each contribute a total of 60 over rounds 1 to 10. Hence, at the beginning of round 11 there are three rich participants with $OF_{i,11} = 140$ and three poor participants with $OF_{i,11} = 100$.

For all treatments it holds true that the group account contains $C_{10} = 240$ at the end of round 10. In order to reach the threshold, the sum of contributions in the 10 active rounds (rounds 11 to 20) has to equal 240. For example, if each participant contributes 4 in each active round, the threshold is reached exactly.

The chosen method of implementing heterogeneity reflects the idea of past wealth inheritance (Tavoni et al., 2011). Besides, there are two further reasons for our decision to start the game with 10 passive rounds. First, due to the contributions in the passive rounds, the group account contains 240 at the beginning of round 11. This ensures that there is no way that the group account can become negative. Even if all participants within a group decide to choose $c_{i,t} = -3$ in all active rounds, the group account will total 60. Second, in the treatments with take options, the simulated contributions in rounds 1 to 10 contain negative contributions. Since all participants are informed about all individual contributions in all rounds, the occurrence of negative contributions might raise participants' awareness to the possibility of negative contributions. We think that both design choices - the impossibility of negative values in the group account and the occurrence of negative contributions - increase the chances of participants choosing negative contributions in the active rounds.

Treatments: In total we conduct six treatments (Table 7.1) which differ according to the distribution of the operating fund at the beginning of round 11 (HOM vs. HET), the actions available to participants within each round (GIVE vs. GIVE-TAKE), and the loss rate *p*. For each treatment, the number at the end corresponds to the loss rate *p*. The potential increase in the operating fund due to the existence of the take options is counteracted by *p*, which is higher in the GT-treatments. We discuss the resulting effects on different equilibria below in section 7.3.3.

	HOM		HET	
	GIVE	GIVE-TAKE	GIVE	GIVE-TAKE
p low	G-HOM55	GT-HOM80	G-HET55	GT-HET80
p high	-	-	G-HET70	GT-HET95

Table 7.1: Treatment characteristics.

Procedures: We conducted 15 experimental sessions at the WISO Experimental Lab in Hamburg in July, October, and December 2020. We had three treatments with 60 participants

in 10 groups, two treatments with 66 participants in 11 groups, and one treatment with 48 participants in 8 groups. Hence, we had a total of 360 participants. Participants were recruited using hroot (Bock et al., 2014) and the experiment was programmed in z-Tree (Fischbacher, 2007). Once all participants of a session had taken their seats in the laboratory, the instructions were read aloud by the experimenter. In the instructions, detailed screenshots of the experiment were shown as well as various examples of possible courses of the experiment and the resulting payoffs. Test questions at the beginning of the experiment ensured that the participants had understood everything correctly. The experiment was followed by questionnaires on environmental attitudes, risk preferences, and demographic characteristics of the participants. To measure participants' environmental attitudes and concerns, we used the revised New Environmental Paradigm (NEP) scale by Dunlap et al. (2000) in the German translation of Schleyer-Lindemann et al. (2018). To measure participants' risk preference, we relied on the well validated question from Dohmen et al. (2011). The exchange rate was 1 = € 0.05 and participants, on average, earned € 14.13 (including a € 5.00 show-up fee). The duration of each session was approximately 60 minutes.

7.3.2 Equilibria

Our variant of the CRSD has multiple equilibria. In the following, we focus on equilibria of the subgame starting at the beginning of round 11. There is a *bad equilibrium*, in which all participants choose the lowest possible contribution in each round, i.e., $c_{i,t} = 0$ in the G-treatments and $c_{i,t} = -3$ in the GT-treatments (and for $t \geq 11$). In these cases, the threshold will not be reached. In addition, all strategy profiles in which the threshold is reached exactly (i.e. $S = 480$) and no participant's payoff is below the payoff of the corresponding bad equilibrium are equilibria. Call these the *good equilibria*. Table 7.2 shows the expected payoffs for bad equilibria and two different types of good equilibria.

Treatment	Participant type	I. Bad eq.	II. Equal contributions eq. (good eq.)	III. Equal payoffs eq. (good eq.)	Diff.: II. – I.	Diff.: III. – I.
G-HET55	Poor	154	180	200	26	46
	Rich	194	220	200	26	6
GT-HET80	Poor	154	180	200	26	46
	Rich	194	220	200	26	6
G-HET70	Poor	136	180	200	44	64
	Rich	176	220	200	44	24
GT-HET95	Poor	136	180	200	44	64
	Rich	176	220	200	44	24
G-HOM55	-	174	200	200	26	26
GT-HOM80	-	174	200	200	26	26

Table 7.2: Expected payoffs from bad and good equilibria, and differences between bad and good equilibria over treatments and participant types.

Out of all *good equilibria*, two stand out because in these equilibria, certain fairness norms are salient. First, there is an equilibrium in which each participant contributes 40. Call this the *equal contributions equilibrium*. Second, there is an equilibrium in which the threshold is reached and the corresponding contributions result in perfect equality payoffs. Call this the *equal payoffs equilibrium*. In the HOM-treatments, these two equilibria coincide.

The *equal contributions equilibrium* is characterized by $\sum_{t=11}^{20} c_{i,t} = 40$ (on average $c_{i,t} = 4$ in rounds 11 to 20) for all participants with payoffs of $\pi_{i,P} = 180$ for poor and $\pi_{i,R} = 220$ for rich participants. Over all 20 rounds, each poor participant in this equilibrium contributes a total of 100 to the group account while rich participants contribute 60. Consequently, when reaching the set threshold, 300/480=62.5% of the group contributions are provided by poor participants and 180/480=37.5% by rich participants.

The *equal payoffs equilibrium* is characterized by $\sum_{t=11}^{20} c_{i,t} = 20$ (on average $c_{i,t} = 2$ in rounds 11 to 20) for poor participants and $\sum_{t=11}^{20} c_{i,t} = 60$ (on average $c_{i,t} = 6$ in rounds 11 to 20) for rich participants. The corresponding payoffs are $\pi_i = 200$ for all participants. In this equilibrium (and considering all rounds 1 to 20), 50% of the group contributions are provided by poor participants and 50% by rich participants.

In the *bad equilibrium*, the expected payoff is $OF_{i,11} + (1-p)e_i$ for the G-treatments and $OF_{i,11} + 30 + (1-p)e_i$ for the GT-treatments. That is, with the take option being available, participants with the same endowment and operating fund will receive a higher payoff if p would be identical across treatments. Also, note that the expected payoffs in all good equilibria are independent of p. Hence, if p would be identical across treatments, the difference between the bad and a specific good equilibrium would be higher for the G-treatments, compared to the corresponding GT-treatments. Consequently, coordination on the good

equilibrium might be easier in the GT-treatments. In order to have a constant bad equilibrium payoff across treatments, we chose higher values for the loss rate *p* in the GT-treatments.

Table 7.2 shows that the differences between good and bad equilibria for the G-treatments and GT-treatments within a corresponding treatment pair are always identical if we consider the good equilibrium to be the *equal contributions equilibrium*. When considering the *equal payoffs equilibrium*, the differences are only identical for poor and rich participants within a corresponding treatment pair.

7.3.3 Hypotheses

Our first set of hypotheses focuses on the effect of the take option. As described above, Krupka and Weber (2013) showed that behavior depends on norms, which define the "social appropriateness" of different actions. Between G- and GT-treatments, behavior might differ because the different action sets might lead to different norms. The lowest possible contribution is 0 in the G-treatments and -3 in the GT-treatments. Consequently, contributing 0 might be considered more acceptable in the GT-treatments, compared to the G-treatments. One would generally expect that small contributions are more acceptable when the take option is present. These differences in "social appropriateness" might increase the frequency of small contributions. This increase might in turn affect participants' beliefs, so that they expect other group members to make smaller contributions in subsequent rounds.

In addition, in the GT-treatments it is possible that even if the threshold was already reached before the last round, withdrawals from the group account in the last round are so large that the threshold cannot be reached in the final round. This is not possible in treatments where the action sets only include give options. Once the threshold is reached in these treatments, it cannot be reduced in a subsequent round.

Taken together, these arguments imply that participants are less likely to reach the threshold in the GT-treatments, as compared to the G-treatments. Let the success rate be the share of groups whose aggregate contributions reach or exceed the threshold. We then test our prediction by comparing the success rates between pairs of treatments which differ only in the availability of the take option and the loss rate *p*. As mentioned above, the difference in loss rates ensures that the payoffs from the three equilibria are identical across both treatments (see also Table 7.2).

Then, our first three hypotheses are:

Hypothesis 1a: The success rate is higher in G-HET70, compared to GT-HET95.

Hypothesis 1b: The success rate is higher in G-HET55, compared to GT-HET80.

Hypothesis 1c: The success rate is higher in G-HOM55, compared to GT-HOM80.

Our next set of hypotheses is based on the results of previous studies, which show that a higher loss rate leads to a higher success rate. As explained above, the higher the loss rate, the larger the difference between the expected payoffs in the bad equilibrium and the good equilibrium. Put bluntly, if p is higher, participants can gain more if the groups' contributions reach or exceed the threshold. We test this prediction by comparing the success rates between pairs of treatments which differ only in the loss rate p. Thus, our next hypotheses are:

Hypothesis 2a: The success rate is higher in GT-HET95 compared to GT-HET80.

Hypothesis 2b: The success rate is higher in G-HET70 compared to G-HET55.

Our last hypotheses focus on the effect of heterogeneity, or, more precisely, the distribution of operating funds within a group after round 10.

As shown in Table 7.2, in the HOM-treatments, the *equal payoff equilibrium* and the *equal contributions equilibrium* are identical. This is not the case in the HET-treatments. Both good equilibria correspond to different fairness principles. Hence, if participants hold different fairness principles, this will make coordination more difficult in the HET-treatments. We therefore expect that coordination to a good equilibrium (and reaching the threshold) is more likely in the HOM-treatments, compared to the HET-treatments. We test this prediction by comparing the success rates between pairs of treatments which differ only in the distribution of the operating fund at the beginning of round 11. Thus, our hypotheses are:

Hypothesis 3a: The success rate is higher in GT-HOM80 compared to GT-HET80.

Hypothesis 3b: The success rate is higher in G-HOM55 compared to G-HET55.

7.4 Results

This section presents the results. In section 7.4.1, we start by describing the general results and successively examine the hypotheses derived above. We then take a closer look at individual contributions and the take option (section 7.4.2), before we examine how burden sharing affects group success in the HET-treatments (section 7.4.3). We conclude the results section by

analyzing how a group member's behavior in early active rounds can predict the group's success (section 7.4.4). Since we find no effect of demographic variables or personal attitudes, we delegate the corresponding analysis to Appendix A.3 (section 7.6.3).

7.4.1 Main results

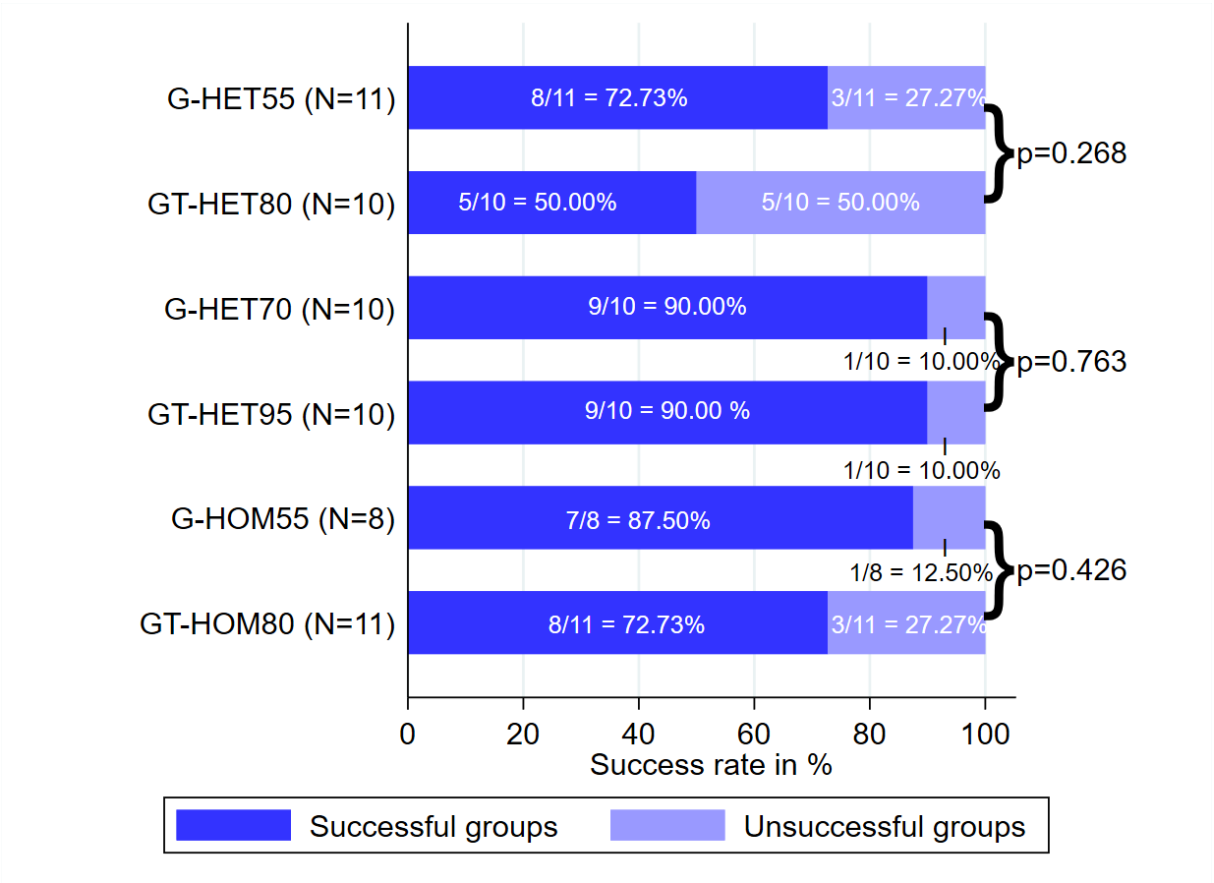


Figure 7.2: Success rates for each treatment.

Notes. *p*-values from one-sided Fisher's exact tests for all pairwise comparisons (see Hypothesis 1).

Figure 7.2 shows the success rates for each treatment.[43] In order to analyze the effect of the take option, we compare success rates between pairs of treatments which differ only in the availability of the take option and the loss rate p. That is, we compare success rates between treatments G-HET55 and GT-HET80, G-HET70 and GT-HET95, and between G-HOM55 and GT-HOM80. In those treatments with low loss rates, success rates are lower when the take option is available (72.8% in G-HET55 vs. 50% in GT-HET80 and 87.5% in G-HOM55 vs. 72.8% in GT-HOM80). In both treatments with high loss rates (G-HET70 and GT-HET95) success rates are identical at 90%.

[43] Mean values and standard deviations for successful and unsuccessful groups for each treatment can be found in Table 7.5 (Appendix A.2, section 7.6.2).

To test whether the results described above are statistically significant, we compare success rates using one-sided Fisher's exact tests. For all pairwise comparisons, the negative effect of the take option is not statistically significant. The resulting *p*-values are 0.268 for G-HET55 vs. GT-HET80, 0.763 for G-HET70 vs. GT-HET95, and 0.426 for G-HOM55 vs. GT-HOM80. Hence, we have to reject Hypotheses 1a, 1b, and 1c.

Result 1: The existence of a take option does not adversely affect success rates.

The success rates are highest (90%) in G-HET70 and GT-HET95, i.e. those treatments in which the loss rates are high and participants have heterogeneous operating funds after round 10. Success rates are lower in the HET-treatments with lower loss rates (72.7% in G-HET55 and 50% in GT-HET80). Apparently, a higher loss rate makes success more likely, which is in line with Milinski et al. (2008). The difference in success rates between the GT-treatments is statistically significant (GT-HET80 vs. GT-HET95, p=0.070), while the difference between the G-treatments is not (G-HET55 vs. G-HET70, p=0.331). Hence, we find support for Hypothesis 2a, but have to reject Hypothesis 2b.

Result 2: A higher loss rate leads to a higher success rate, but only when the take option is available.

Comparing the treatments which differ only with respect to the distribution of the operating fund at the beginning of round 11, we see that success rates are higher in the HOM-treatments (87.5% in G-HOM55 and 72.7% in GT-HOM80), compared to the corresponding HET-treatments (72.7% in G-HET55 and 50% in GT-HET80). This indicates a negative effect of heterogeneity and is in line with the results in Tavoni et al. (2011) and Burton-Chellew et al. (2013). However, the negative effect of heterogeneity is statistically not significant (p=0.268 for GT-HOM80 vs. GT-HET80 and p=0.426 for G-HOM55 vs. G-HET55), forcing us to reject Hypotheses 3a and 3b.

Result 3: Heterogeneity in the operating funds (after round 10) does not negatively affect the success rates.

7.4.2 Individual contributions and the take option

In the previous section, we looked at the results aggregated at the group level. In this section, we look at individual contributions for all active rounds (11-20), for the first active round (11), and additionally take a closer look at the take option.

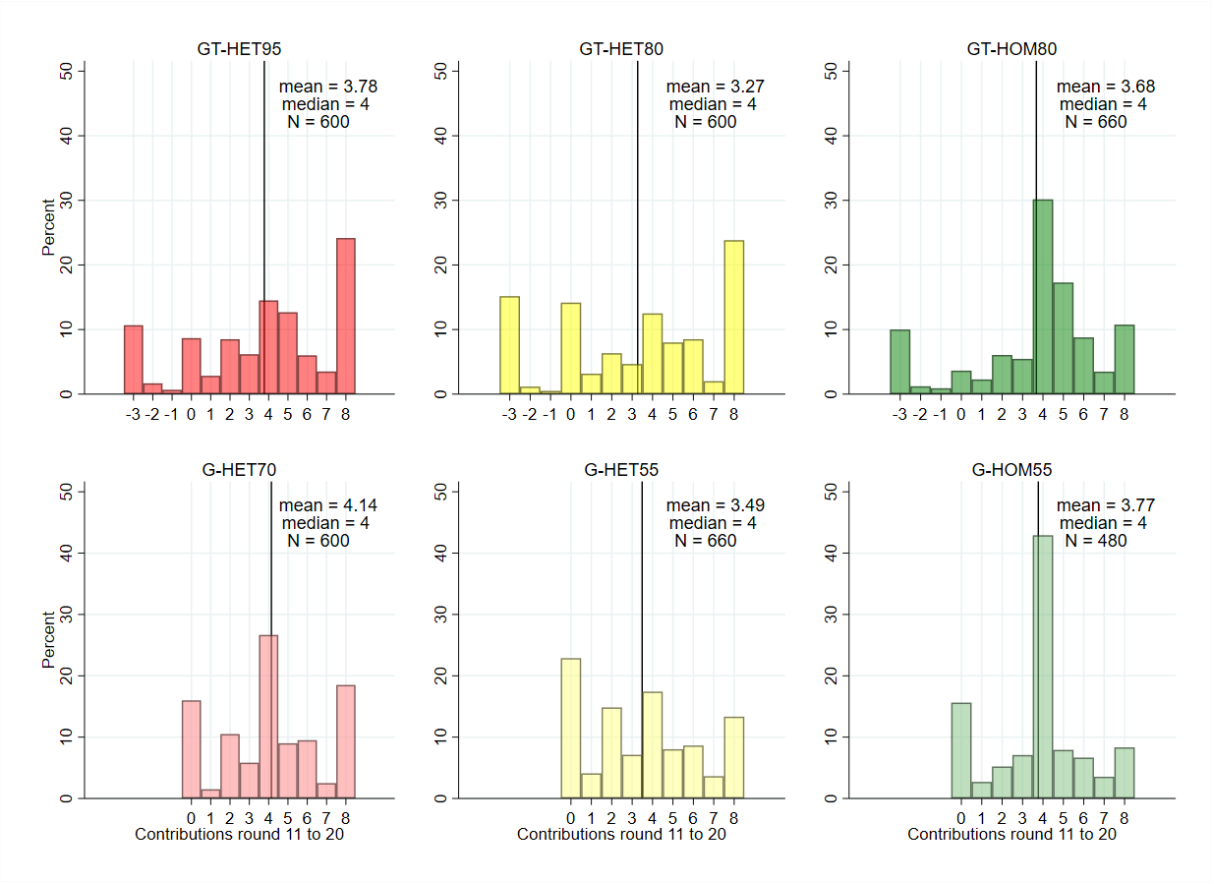


Figure 7.3: Distribution of contributions in rounds 11 to 20 by treatment.

Notes. Vertical lines indicate mean values.

Figure 7.3 shows the distribution of contributions by treatment for all active rounds, as well as the mean and median values. For all treatments, the mean is between 3.27 and 4.14 and the median is 4. In the GT-treatments, we see that a take option ($c_{i,t} < 0$) was chosen in at least ten percent of all cases.

Focusing on mean contributions, it can be seen that means are smaller when the take option is available. In Figure 7.3, each GT-treatment in the top row has a smaller mean than the corresponding G-treatment in the bottom row. A pairwise comparison of means between treatments which differ only with respect to the loss rate (i.e., the red and yellow treatments in Figure 7.3) suggests that a higher loss rate is associated with higher mean contributions. A pairwise comparison of means between treatments which differ only with respect to the distribution of the operating fund at the beginning of round 11 (i.e., the yellow and green

treatments in Figure 7.3) shows that heterogeneity is associated with smaller mean contributions.[44]

Figure 7.4: Distribution of contributions in round 11 by treatment.

Notes. Vertical lines indicate mean values.

Figure 7.4 shows the distribution of contributions in round 11 by treatment, as well as the mean and median values. Again, mean values are in a narrow range between 3.78 and 4.97, and the median contribution is 4 in all treatments. Since round 11 contributions are independent observations, we can conduct statistical tests using individual contributions in round 11 as observations. Due to the non-normal distribution of contributions, we use non-parametric two-sided Mann-Whitney-tests. In all three G-treatments the mean is higher, compared to the corresponding GT-treatments. However, there are no significant differences between the corresponding treatments (all p-values>0.05). In the HET-treatments, a higher loss rate is associated with higher mean contributions in round 11. Yet, the difference is only significant between G-HET70 and G-HET55 (N=126, p-value=0.031). Although mean contributions in

[44] We do not conduct any statistical tests comparing means or medians (of all contributions between rounds 11 and 20) because within each group, contributions are dependent.

both HOM-treatments are higher than in the corresponding HET-treatments, these differences are not significant (both *p*-values>0.05).

Treatment	Type	Successful groups	Unsuccessful groups	All groups
GT-HET95	rich	4 (1.48%)	19 (63.33%)	23 (7.67%)
	poor	22 (8.15%)	19 (63.33%)	41 (13.67%)
GT-HET80	rich	1 (0.67%)	42 (28.00%)	43 (14.33%)
	poor	7 (4.67%)	41 (27.33%)	48 (16.00%)
GT-HOM80		21 (4.38%)	45 (25.00%)	66 (10.00%)
All three		55 (4.17%)	166 (30.74%)	221 (11.88%)

Table 7.3: Frequencies and shares of take3-decisions in GT-treatments.

The existence of the take option is the central characteristic that distinguishes our version of the CRSD from existing experimental investigations of the CRSD. In the remainder of this section, we take a closer look at those treatments in which the take option is available. To do so, we focus on the most extreme take option, namely the decision to take 3 out of the group account ($c_{i,t} = -3$, henceforth take3-decisions). Figure 7.3 shows that the vast majority of all chosen take options were take3-decisions (82.05% in GT-HET95, 90.10% in GT-HET80, 82.50% in GT-HOM80).

Table 7.3 shows the frequency and the relative share of take3-decisions among all contribution decisions. Three points are worth highlighting. First, participants are willing to use the take3-decision. Over all GT-treatments, 11.88 % of all contributions (i.e. both take and give) are take3- decisions (see also Figure 7.3). Second, in the HET-treatments, the majority of take3-decisions are made by poor participants, although this effect is driven by poor participants in successful groups. In unsuccessful groups, both rich and poor participants chose the take3-decision equally often. And third, across all GT-treatments, there were four groups which already reached the threshold at the end of round 19. In these groups, no single participant chose a take option in round 20 (see Table 7.6 in Appendix A.2, section 7.6.2).

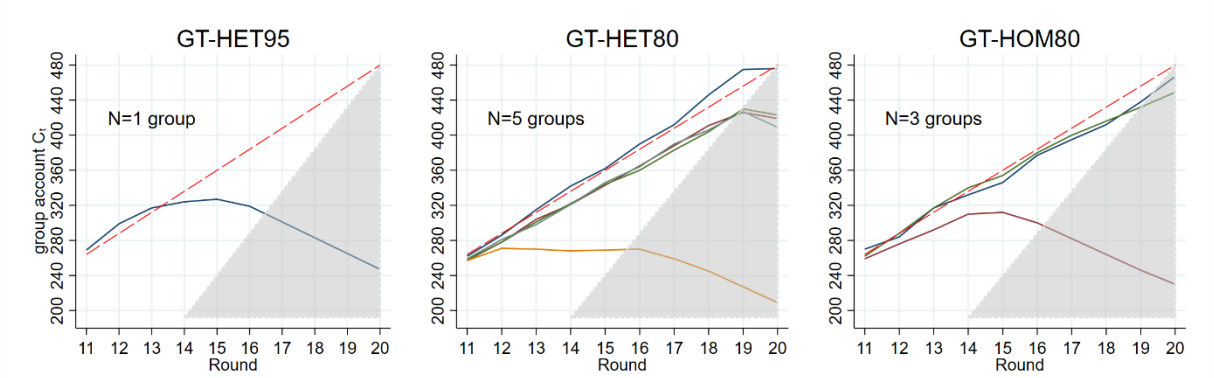


Figure 7.5: Evolution of group accounts in unsuccessful groups over GT-treatments.

Notes. The individual curves show the evolution of group accounts for each unsuccessful group within the respective GT-treatment. The red dashed line shows the evolution of the group account for the case that a total amount of 24 is contributed in each round and the threshold of 480 is thus exactly reached.

Since the take option is rarely used in successful groups, we focus on unsuccessful groups in the rest of this section. Figure 7.5 depicts the evolution of the group account (C_t) over all active rounds.[45] The gray area in the lower right contains all possible values of C_t, for which it is impossible to reach the threshold.

We see that in each of the three GT-treatments, there is one group for which the group account decreases before the gray area is reached. In treatments GT-HET95 (blue line) and GT-HOM80 (dark red line), the decrease starts at round 15, and in treatment GT-HET80 (yellow line), it starts even earlier. From the point where the blue (GT-HET95), yellow (GT-HET80), and dark red (GT-HOM80) lines cross the gray area, it can be seen that the impossibility to reach the threshold only became apparent in rounds 16 or 17.

For all other unsuccessful groups, the value of the group account increased until rounds 19 or 20. In treatment GT-HET80, there is one group with $C_{19} = 475$. This group failed to reach the threshold because aggregate contributions in round 20 were only 1 (individual contributions: -3,-3,+1,+2,+2,+2). For the other four unsuccessful groups in GT-HET80, the threshold could not be reached any more ($C_{19} < 432$) and the take3-decision was selected by 22 out of 24 participants (91.67%)[46]. In treatment GT-HOM80, the two remaining unsuccessful groups could still reach the threshold in round 20 ($C_{19} = 432$ and $C_{19} = 438$), but failed because round 20 contributions were 17 and 29 respectively.

[45] In Appendix A.2 (section 7.6.2), Figure 7.8 shows the respective graphs for all unsuccessful groups in G-treatments and Figure 7.9 for all successful groups in GT- and G-treatments.

[46] The other two participants in these groups chose contributions of 8.

Result 4: The majority of chosen take options is the most extreme take3-decision. In some groups, the group account was decreasing even though the threshold could still be reached. Once it was obvious that reaching the threshold was no longer possible, almost all participants selected the take3-decision.

7.4.3 Burden sharing in HET-treatments

In this section, we focus on HET-treatments and explore how burden sharing between rich and poor participants differs between successful and unsuccessful groups. After the 10 passive rounds (1-10), each group had a group account totaling 240. Thus, at the end of the passive rounds, the three poor participants had contributed a total of 180 (a relative share of 75%) and the three rich participants a total of 60 (a relative share of 25%).

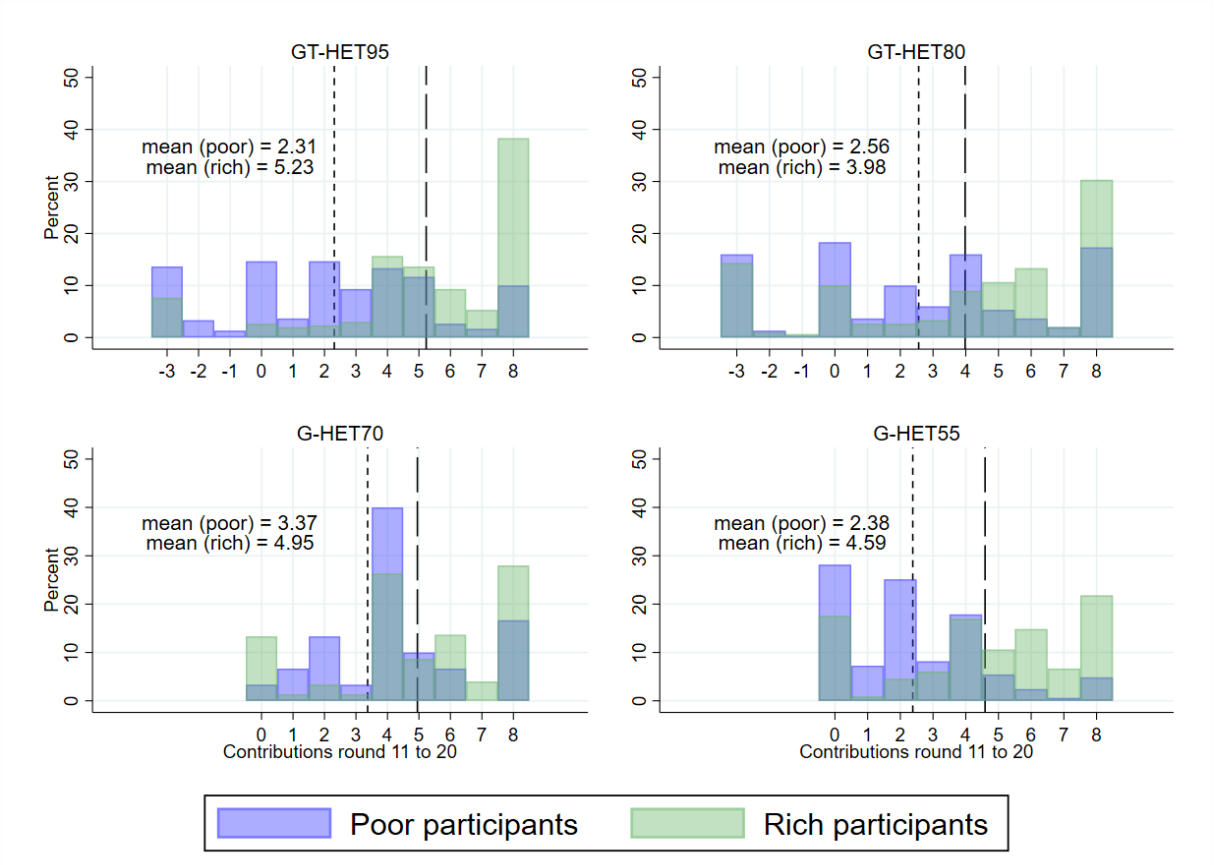


Figure 7.6: Distribution of contributions in rounds 11-20 by participant type and HET-treatment.

Notes. Vertical short (long) dashed lines indicate mean values for poor (rich) players.

Figure 7.6 shows the distribution of contributions and mean values for rich and poor participants for all active rounds. On average, poor participants contributed between 2.31 and 3.37 while

rich participants contributed between 3.98 and 5.23. Besides the finding that the highest contribution of 8 is chosen more often by rich participants, no clear pattern can be observed.[47]

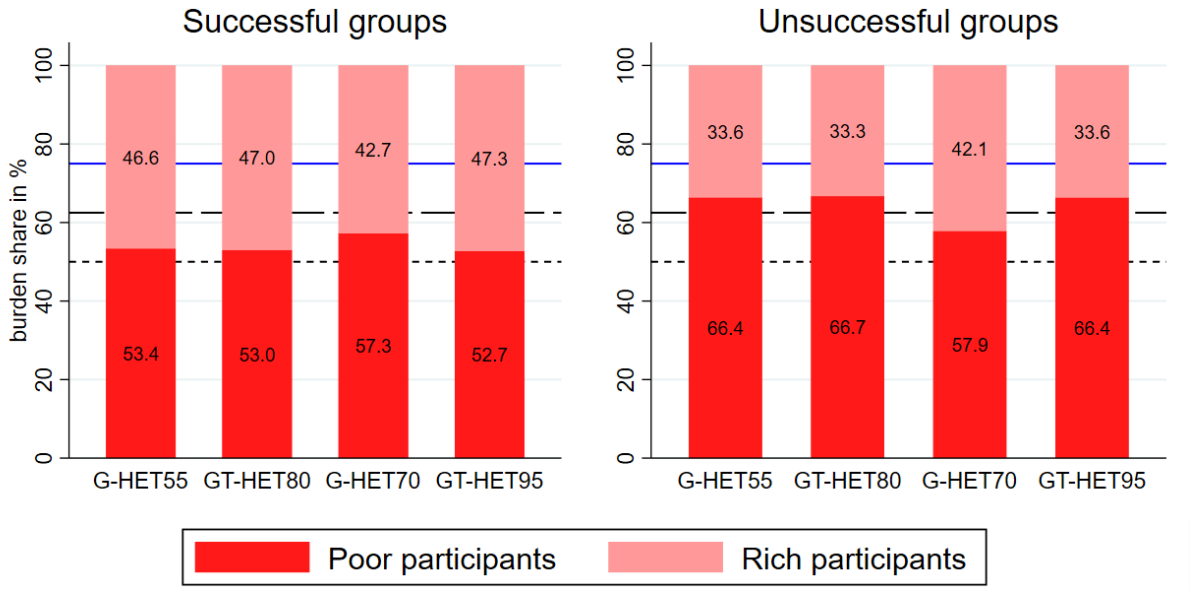


Figure 7.7: Burden sharing between poor and rich participants in successful and unsuccessful groups.

Notes. The values within the bars indicate the average percentage of the group account contributed by poor and rich participants over all 20 rounds. The solid blue line marks the distribution of contributions after round 10. The long dashed lines correspond to the relative shares which would result from the equal-contributions equilibrium. The short dashed lines correspond to the relative shares which would result from the equal-payoffs equilibrium.

Figure 7.7 shows the relative shares of total contributions over all 20 rounds. The relative share at the beginning of the active rounds is indicated by the solid blue line. Although rich participants contributed more than poor participants in the 10 active rounds, the difference in contributions over rounds 11 to 20 was not sufficient to establish an equal burden sharing between poor and rich participants. Over all 20 rounds, rich participants contributed, on average, less than poor participants. This holds true regardless of treatment type and whether the threshold was reached or not. Within each treatment, however, the rich participants' burden share is higher for successful groups, compared to unsuccessful groups.

Result 5: In all groups in the HET-treatments, the rich compensated the poor over rounds 11 to 20, but successful groups are characterized by higher compensations.

7.4.4 Predicting a group's success

We saw that in most groups, the take option is primarily used once it is apparent that the threshold cannot be reached any more. In rounds 11 to 13, however, it is always possible to

[47] Higher contributions of rich participants and a higher frequency of contributions of 8 can also be seen in the first active round. Figure 10 in Appendix A.2 (section 7.6.2) shows the distributions and means for round 11.

reach the threshold[48], and yet, we observe that some participants use the take option in these rounds. The early use of the take option might negatively affect participants' beliefs, such that they expect lower contributions in subsequent rounds (see section 7.3.3). If participants condition their own contribution on their belief about others' contributions, this could result in a vicious circle.

Similarly, the burden share could affect participants' beliefs. A low burden share of rich participants could signal rich participants' unwillingness to take a large burden and, hence, might negatively affect beliefs of poor participants.

Although we do not have data on beliefs, we can test these conjectures by running a probit regression in which we include the number of take3-decisions and the burden share of rich participants at the end of round 13 as independent variables. The binary outcome variable indicates whether the threshold was reached (=1) or not (=0). Table 7.4 shows three specifications.

	(1)	(2)	(3)
Group contribution (rounds 11-13)		0.07*** (0.02)	0.05*** (0.02)
HET	-0.31 (0.49)	-0.13 (0.51)	-8.25** (3.53)
p high	0.99** (0.49)	0.46 (0.54)	0.49 (0.59)
Take	-0.39 (0.46)	-0.46 (0.57)	-0.03 (0.66)
Take*take3-decisions (rounds 11-13)			-0.27 (0.28)
HET*burden share rich (rounds 11-13)			24.85** (11.09)
Constant	0.80** (0.33)	-3.89*** (1.25)	-3.16** (1.40)
N	60	60	60
Wald Chi^2	4.56	15.28	32.07
Pseudo R^2	0.08	0.32	0.42

Table 7.4: Probit regressions of group success on group contribution behavior in rounds 11-13.

Notes. Standard errors clustered by group in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Model (1) includes only indicator variables for the heterogeneity of the operating fund (HET), the loss rate (*p* high), and the existence of the take option (Take). As hypothesized, both the

[48] At the beginning of round 11, the group account contains 240. If every participant contributes -3, the group account decreases by 18 per round, so that at the end of round 14, the group account contains 240-4*18=168. In the remaining six rounds, the maximum contribution by all six group members is 6*6*8=288. Since 168+288=456, the threshold (480) would not be reached. If all group members contribute -3 only in rounds 11, 12 and 13 and contribute 8 thereafter, the group account would contain 522 at the end of round 20 and the threshold would be reached.

existence of the take option and heterogeneity have a negative effect, although these effects are insignificant. Only the loss rate has a significant positive effect. In model (2), we add the sum of contributions over rounds 11-13 (Group contribution) as independent variable. While the loss rate effect is no longer significant, group contributions have a positive significant effect, which can be linked to the correlation between loss rates and contributions over rounds 11-13.

In model (3), we add the frequency of take3-decisions (only for GT-treatments) and the burden share of rich participants after round 13 (only for the HET-treatments). The frequency of take3-decisions has no significant effect, although this might be due to the multicollinearity between take3-decisions and group contributions. Despite heterogeneity now showing a significant negative effect, the interaction with rich participants' burden share is positive and significant. That is, in the HET-treatments, groups in which the rich signal their willingness to contribute (comparatively more) in rounds 11 to 13 are more likely to reach the threshold.

Result 6: In the HET-treatments, groups are more likely to reach the threshold when rich participants signal their willingness to contribute early on.

7.5 Discussion and conclusion

The possibility of exploiting public goods or the threat of exploitation characterizes numerous conflicts in the world outside the laboratory. By analyzing the effects of a take option, our experiment considers an important aspect of climate change-related coordination problems that has not been covered by previous experimental literature on the so-called collective-risk social dilemma (CRSD). Ignoring the existence of a take option restricts not only the external validity of the existing studies on the CRSD, but might also bias the resulting policy implications. By conducting a laboratory experiment which extends the previous literature on the CRSD, we aim to fill this research gap.

The CRSD used in our experiment is a threshold PGG in which groups of six participants interact over ten active rounds. Within each round, participants simultaneously contribute to a group account and the public good is then only provided if the sum of contributions at the end of the game exceeds a commonly known threshold. In the G-treatments, participants can only choose non-negative contributions, while in the GT-treatments participants can choose their contributions from a larger action set which also includes negative contributions.

Our main research question was to investigate whether achieving a threshold in such a CRSD is more difficult in GT-treatments than in G-treatments. Our results show that, with a low loss

rate, success rates are lower when the take option is available. However, these effects are statistically not significant. For treatments with high loss rates, we find no difference in success rates. This is in line with the results of Khadjavi and Lange (2015) who also do not find significant differences between contributions in G- and GT-treatments in repeated linear PGGs.

Besides the effect of the take option, we investigated how changes in the loss rate and heterogeneity in the distribution of the operating fund (the pot of money out of which contributions are paid) affect cooperation. Regarding the effect of the loss rate, we find that a higher loss rate leads to more groups reaching the threshold. This effect is statistically significant but only for the GT-treatments. For heterogeneity, we find that success rates are lower for groups with heterogeneous operating funds, but again, the effect is not statistically significant.

Although the participants in our experiment have wider action sets than in similar studies, we observe that in five of the six treatments, about half of the participants already choose contributions that correspond to one of the described good equilibrium paths in the first active round. We further find that a group's total contributions in the first three active rounds have a high predictive power for group success. In contrast, when controlling for the value of the group account at the end of the first three active rounds, the number of take3-decisions in the first three active round has no predictive power for the group's success.

For HET-treatments, we find that the burden share of rich participants (i.e., the share of total contributions made by rich participants) is a key driver of group success, which confirms one key finding of Tavoni et al. (2011).

As shown in Figure 7.5, in each GT-treatment, there is one group whose group account is already decreasing, although the threshold could still be reached. In addition, Figure 8 (Appendix A.2, section 7.6.2) shows that in all HET-treatments, both poor and rich participants already choose the smallest possible contribution (0 or -3) in the first round. Figure 7.4 in section 7.4.2 shows that similar decisions can also be observed in the HOM-treatments. This could indicate that in all six treatments, several participants were pessimistic about reaching the threshold already at the beginning of the active rounds. Alternatively, it could suggest that these participants chose the smallest possible contributions in order to force other group members to make higher contributions.

Although the direction of the effects is mostly as expected, results are predominantly not statistically significant. This implies that we cannot make a conclusive statement regarding the

take option effect. Most of our results are based on one-sided Fisher's exact tests which we use for the pairwise comparisons of treatments which differ only with respect to one characteristic. This is a very conservative approach which will detect significant effects only if they are large enough. In order to get more conclusive results, further experiments could use a larger number of groups per treatment.

Our results contribute to a better understanding of the cooperation dynamics in climate change experiments. We hope that they may motivate future research to investigate which institutions (see e.g., Dannerberg and Gallier, 2019; Chaudhuri, 2011; Gürerk et al., 2006) are appropriate to effectively promote cooperation and prevent the exploitation of (global) public goods. Furthermore, future studies could use similar experimental settings to investigate whether cooperation is hampered when a take option is available only for some group members or when the range of a take option varies between group members. The effect of a take option could also be examined in CRSD versions with uncertain or ambiguous thresholds (Brown and Kroll, 2017; Dannenberg et al., 2015) in intergenerational PGGs on climate change (e.g., Böhm et al., 2020; Lohse and Waichman, 2020).

Acknowledgments

We thank Philipp Balkwitz and the team of the WISO Experimental Lab in Hamburg for their support in conducting the experiment. For helpful feedback and comments, we thank Merindah Loessl, Svenja Mohr, and the participants of the 2020 Economic Science Association (ESA) Global Online Around-the-Clock Meetings.

7.6 Appendix

7.6.1 A.1 Instructions (translated from German)

Welcome to the experiment and thank you for participating

General Information

Please read these instructions carefully. Do not talk to the other participants during the whole experiment. If you have any questions, please contact the experimenter silently by raising your hand. We will then come to you and answer your questions. Compliance with this rule is very important. Otherwise, the results of this experiment lose their scientific value.

Please take sufficient time to read the explanations when making your decisions. You cannot influence the duration of the experiment by making a quick decision, as you always have to wait for the other participants. The experiment is completely anonymous. During the experiment or afterwards, you will not be informed with whom you have interacted. The other participants will not, neither during the experiment nor afterwards, receive information on your decisions and final earnings.

For your participation in this experiment you will receive a fixed payment of € 5.00 and an **additional variable payment.** At the end of the experiment, a questionnaire follows. Afterwards, all participants of the experiment are called one after the other and receive their payment. The payout is private, anonymous, and in cash. The duration of the experiment is **approximately 60 minutes.**

The Experiment

Today's experiment involves 30 people who are divided into **five groups of six group members** each. The group allocation is carried out randomly. During the experiment you will only interact with the other members of your group of six. Within a group, each group member is randomly assigned a pseudonym which is displayed at the top right of the screen during the experiment. The pseudonyms correspond to the names of moons and dwarf planets from our solar system. Thus, you can follow the decisions of the group members while anonymity is guaranteed.
All amounts in the experiment are displayed in the currency Taler. For your final compensation, the exchange rate is 20 Taler = € 1 or 1 Taler = € 0.05.

The starting point: Each group member has an active account of 160 Taler and a passive account of 120 Taler. The experiment consists of a total of 20 rounds. Each group has a common **group account** which contains 0 Taler at the beginning of the first round.

The contribution decision within a round: Within a round, all group members decide simultaneously which amounts they want to contribute to or withdraw from the group account. Each group member can either contribute 8, 7, 6, 5, 4, 3, 2, 1 or 0 Taler from their own active account and add this to the joint group account, or withdraw -1, -2 or -3 Taler from the joint group account and add this to their own active account. The amount of the passive account does not change.

Rounds 1-10: In rounds 1 to 10, the contribution decisions of all group members are made by the computer. That is, a random contribution is selected for each group member and the balances of the active account and group account change accordingly. Positive contributions (8, 7, 6, 5, 4, 3, 2, 1) correspond to payments into the group account, and negative contributions (-1, -2, -3) correspond to withdrawals from the group account. Before initiating round 11, the contribution decisions and account balances from rounds 1 to 10 are displayed in a **results table** (see figure on page 4). For each group member, the results table shows the contribution for each round, the sum of the contributions to the group account, and the total sum in the active account. Furthermore, the average contribution per group member and the total contribution to the group account are displayed.

Rounds 11-20: In rounds 11 to 20, all group members make their own contribution decisions. This means that in each round, all group members decide simultaneously what amount they

want to contribute to or withdraw from the group account. After each round, the results for all group members are displayed in the results table. The results table is visible after each round for a maximum of 60 seconds before the next screen appears. By clicking on the OK button, you can leave the results table screen before the 60 seconds have ended.

Calculation of payments

The variable payment, which you receive in addition to the fixed payment of €5.00, consists of two parts. The first part is the amount that is in your active account after round 20. The second part of the variable payment is the same for all group members and depends on the total balance of the group account:

- If there are **at least 480 Taler** in the group account after round 20, you will receive the **full 120 Taler** from your passive account.
- If there are **less than 480 Taler** in the group account after round 20, 95% (114 Taler) of the passive account will be lost and you will only receive **the remaining 5% (6 Taler)** from your passive account.

Your payment is therefore composed as follows:

Payment = fixed payment + active account after round 20 + passive account

Results table (will be shown after each round from 10 to 20)

The column **Round** shows the respective round.

The columns to the right of it in the top cell indicate the pseudonym (here **Deimos**) of each group member and, in the cells below, the (positive or negative) contribution of the group member to the group account in each individual round.

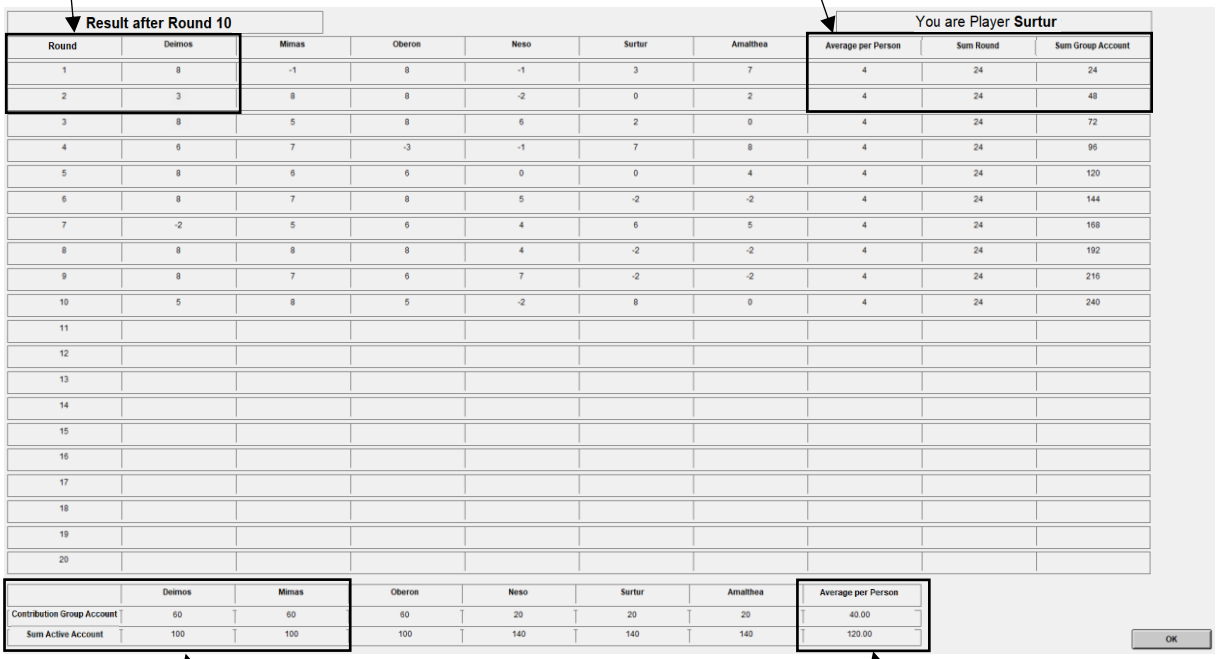
The column **Average per Person** shows the average contribution of the group members to the group account for the respective round.

The column **Sum Round** shows the sum of all group members' contributions to the group account for the respective round.

The column **Sum Group Account** shows the total amount in the group account at the end of the current round.

Round	Deimos
1	8
2	3

Average per Person	Sum Round	Sum Group Account
4	24	24
4	24	48



	Deimos	Mimas
Contribution Group Account	60	60
Sum Active Account	100	100

Average per Person
40.00
120.00

The row **Contribution Group Account** shows each individual group member's contribution to the group account at the end of the current round.

The row **Sum Active Account** shows each group member's current amount in their Active Account.

The column **Average per Person** in the upper cell shows the average contribution to the group account and, in the lower cell, the average amount in the active account by group member at the end of the current round.

Example

There are several possible scenarios over the course of the experiment, two of which are explained as examples. Of course, all other scenarios are also possible over the course of the experiment.

Scenario A:

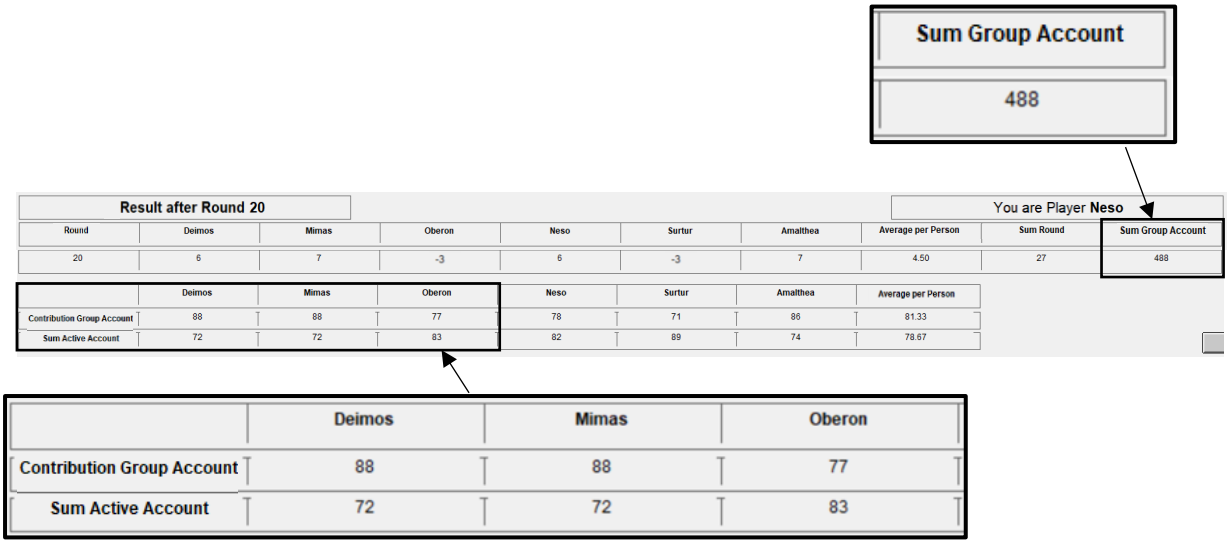
The figure below shows the result after round 20 and, to illustrate the example, we look at the group members Mimas and Oberon.

Mimas has contributed 88 Taler to the group account and still has 72 Taler in the active account. These 72 Taler represent the first part of the variable payment. The group account contains 488 Taler which is more than 480 Taler. Therefore, the 120 Taler of the passive account will be paid out. These 120 Taler represent the second part of the variable payment. The **variable payment** that Mimas receives in addition to the fixed payment of €5.00 is therefore

$$\mathbf{72 + 120 = 192\ Taler\ or\ respectively\ 192 \times €0.05 = €9.60.}$$

Oberon has contributed 77 Taler to the group account and still has 83 Taler in the active account. These 83 Taler represent the first part of the variable payment. The group account contains 488 Taler which is more than 480 Taler. Therefore, the 120 Taler of the passive account will be paid out. These 120 Taler represent the second part of the variable payment. The **variable payment** that Oberon receives in addition to the fixed payment of €5.00 is therefore

$$\mathbf{83 + 120 = 203\ Taler\ or\ respectively\ 203 \times €0.05 = €10.15.}$$



Sum Group Account
488

Result after Round 20 — You are Player **Neso**

Round	Deimos	Mimas	Oberon	Neso	Surtur	Amalthea	Average per Person	Sum Round	Sum Group Account
20	6	7	-3	6	-3	7	4.50	27	488

	Deimos	Mimas	Oberon	Neso	Surtur	Amalthea	Average per Person
Contribution Group Account	88	88	77	78	71	86	81.33
Sum Active Account	72	72	83	82	89	74	78.67

	Deimos	Mimas	Oberon
Contribution Group Account	88	88	77
Sum Active Account	72	72	83

Scenario B:

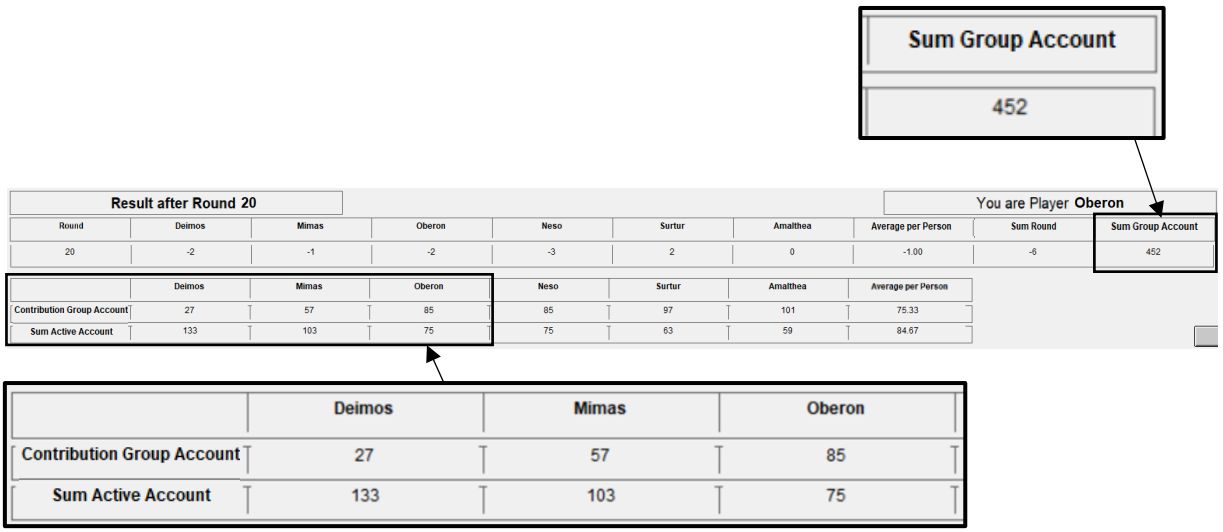
The figure below shows the result after round 20 and, to illustrate the example, we look at the group members Deimos and Mimas.

Deimos has contributed 27 Taler to the group account and still has 133 Taler in the active account. These 133 Taler represent the first part of the variable payment. The group account contains 452 Taler which is less than 480 Taler. Therefore, 95% (114 Taler) of the passive account will be lost and Deimos will only receive the remaining 5% (6 Taler) from the passive account. These 6 Taler represent the second part of the variable payment. The **variable payment** that Deimos receives in addition to the fixed payment of €5.00 is therefore

$$\mathbf{133 + 6 = 139\ Taler\ or\ respectively\ 139 \times €0.05 = €6.95.}$$

Mimas has contributed 57 Taler to the group account and still has 103 Taler in the active account. These 103 Taler represent the first part of the variable payment. The group account contains 452 Taler which is less than 480 Taler. Therefore, 95% (114 Taler) of the passive account will be lost and Deimos will only receive the remaining 5% (6 Taler) from the passive account. These 6 Taler represent the second part of the variable payment. The **variable payment** that Mimas receives in addition to the fixed payment of €5.00 is therefore

$$\mathbf{103 + 6 = 109\ Taler\ or\ respectively\ 109 \times €0.05 = €5.45.}$$



	Deimos	Mimas	Oberon
Contribution Group Account	27	57	85
Sum Active Account	133	103	75

Summary:

- **Each group consists of six group members.**
- **The experiment runs for 20 rounds.**
- **The contribution decisions in rounds 1-10 are made by the computer and in rounds 11-20 by yourself.**
- **Prior to round 1, each group member has 160 Taler in her/his active account and 120 Taler in her/his passive account. Prior to round 1, there are 0 Taler in the group account.**
- **In each round, each group member can either contribute 8, 7, 6, 5, 4, 3, 2, 1 or 0 Taler from their own active account and add this to the joint group account, or withdraw -1, -2 or -3 Taler from the joint group account and add this to their own active account.**
- **If there are at least 480 Taler in the joint group account after round 20, each group member will receive her/his complete passive account of 120 Taler.**
- **If there are less than 480 Taler in the group account after round 20, 95% (114 Taler) of the passive account will be lost and each group member will only receive the remaining 5% (6 Taler) from her/his passive account.**
- **Your payment = fixed payment + active account after round 20 + passive account**.
- **Exchange rate: 20 Taler = €1 or respectively 1 Taler = €0.05.**

If you have any questions, please raise your hand and wait silently until someone comes to you. Please remain silent and do not communicate with the other participants during the whole experiment.

Thank you very much for your participation

7.6.2 A.2 Additional statistical analysis

Treatment	Successful groups	Unsuccessful groups	All groups
GT-HET95	490.89	247.00	466.50
	(7.20)	-	(77.42)
	N=9	N=1	N=10
G-HET70	489.78	477.00	488.50
	(5.85)	-	(6.84)
	N=9	N=1	N=10
GT-HET80	485.40	387.20	436.30
	(3.36)	(102.97)	(86.00)
	N=5	N=5	N=10
G-HET55	488.50	345.00	449.36
	(3.82)	(64.78)	(73.09)
	N=8	N=3	N=11
GT-HOM80	490.00	382.00	460.55
	(2.33)	(131.94)	(77.66)
	N=8	N=3	N=11
G-HOM55	488.71	309.00	466.25
	(3.35)	-	(63.61)
	N=7	N=1	N=8

Table 7.5: Mean group accounts and standard deviations (in parentheses) after round 20 over experimental treatments, ordered by successful and unsuccessful groups.

	$C_{19} \geq 480$ (threshold reached after round 19)			$C_{19} < 480$ (threshold not reached after round 19)		
Treatment	N groups	Mean (sd) contribution in round 20	Minimum contribution in round 20	N groups	Mean contribution in round 20	Minimum contribution in round 20
GT-HET95	2	2.08 (2.68)	0	7	3.24 (2.85)	-3
GT-HET80	2	0.42 (1.44)	0	3	2.94 (2.34)	0
GT-HOM80	0	-	-	8	3.50 (2.43)	-2
G-HET70	3	0.00 (0.00)	0	6	4.67 (2.48)	0
G-HET55	0	-	-	8	3.33 (2.31)	0
G-HOM55	0	-	-	7	4.88 (2.10)	0

Table 7.6: Contribution behavior of successful groups in round 20 ordered by whether the threshold was reached after round 19.

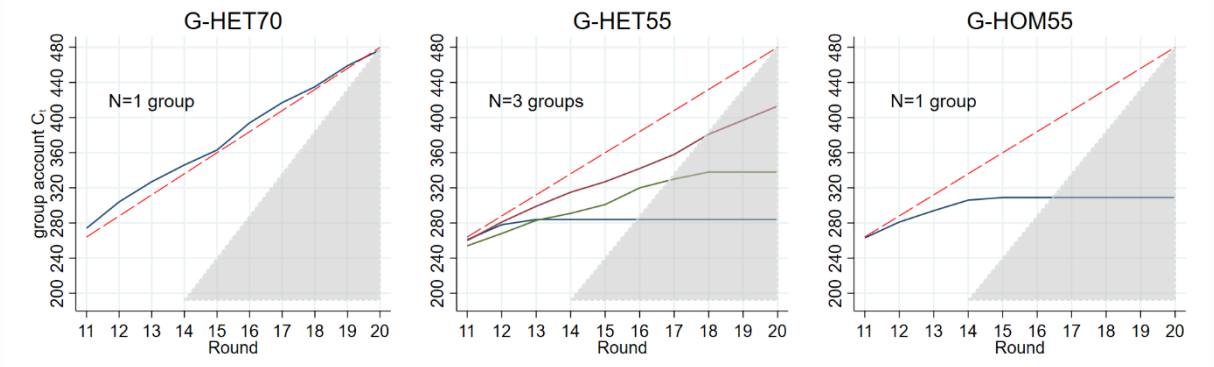


Figure 7.8: Evolution of group accounts in unsuccessful groups over G-treatments.

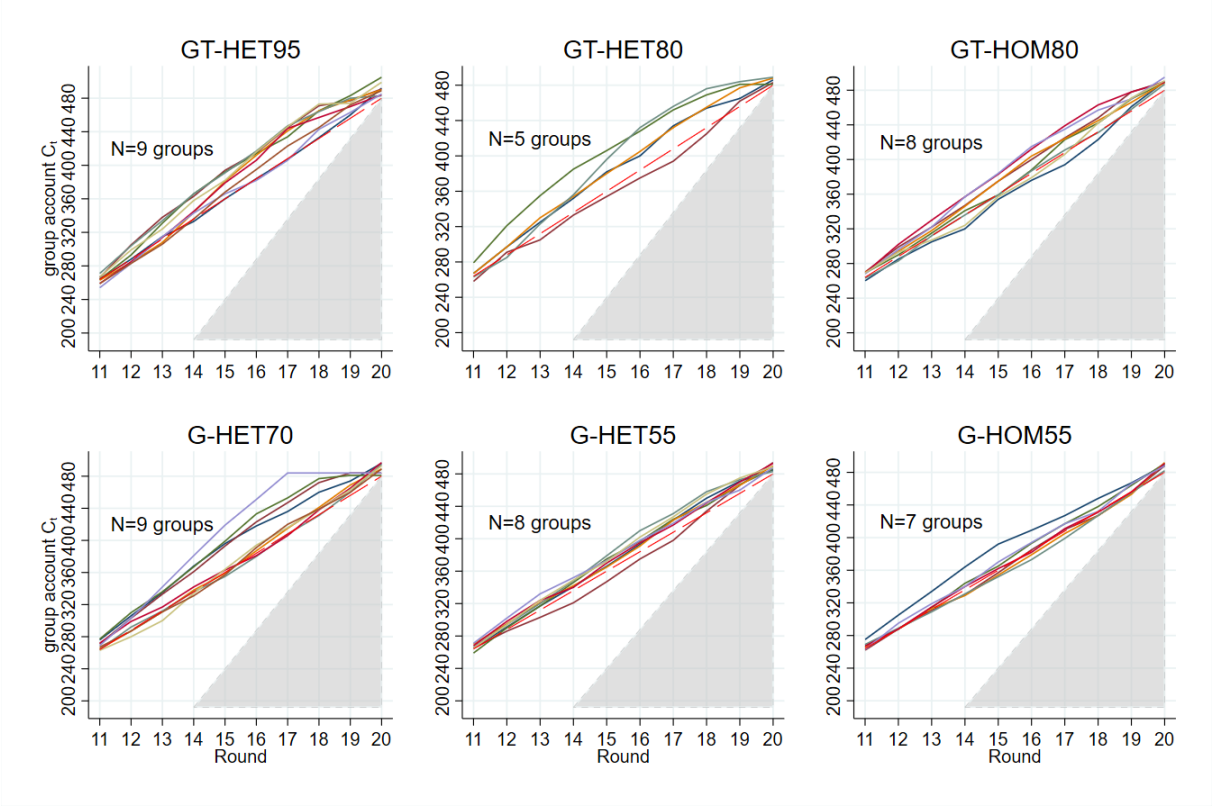


Figure 7.9: Evolution of group accounts in successful groups over GT- and G-treatments.

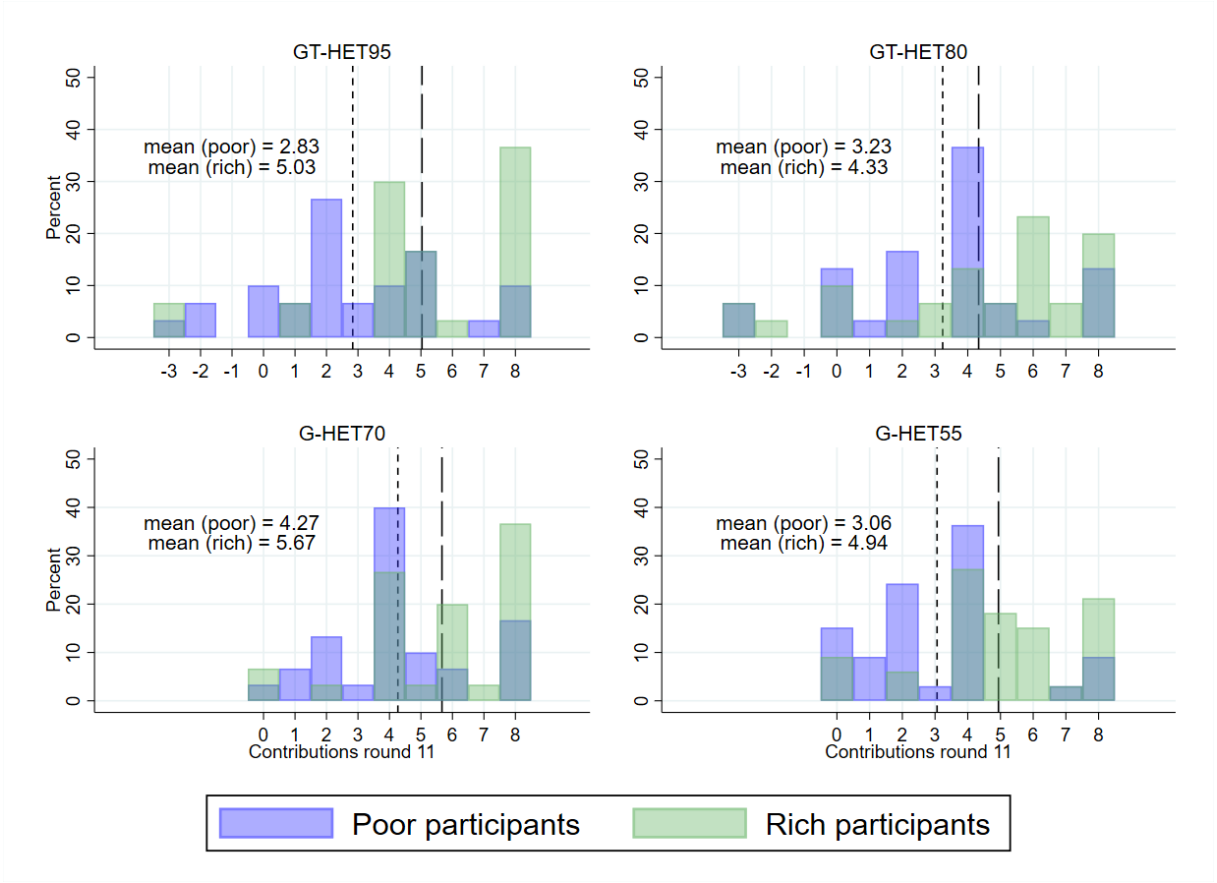


Figure 7.10: Distribution of contributions in round 11 by participant type and HET-treatment.

Notes. Vertical short (long) dashed lines indicate mean values for poor (rich) players.

7.6.3 A.3 The impact of demographics and personal attitudes

In this section, we examine the influence of participants' demographic characteristics and personal attitudes on their contributions in round 11 and on group success. In the OLS regression in Table 7.7, the dependent variable is the individual contribution in round 11 in heterogeneous treatments (model 1) or homogeneous treatments (model 2). Both regression models include independent variables for *Gender* and *Age*, as well as for the previous number of experiments participated in (*Experiments*). *NEP-Index* is the sum of answer values for the NEP scale by Dunlap et al. (2000). *NEP-Index* is calculated from the 11 items[49] that yield the highest Cronbach's alpha value (0.79). *Risk* contains the participants' reported risk propensity. Mean values and standard deviations for each item of the NEP scale and for the risk question can be found in Tables 7.8 and 7.9. Model 1 (heterogeneous treatments) includes a dummy variable which takes a value of 1 for rich players.

As Table 7.7 shows, both demographic characteristics and personal attitudes have no significant effect on contributions in round 11. In the heterogeneous treatments, rich players choose statistically significant higher contributions.

Next, we examine whether demographic characteristics and personal attitudes affect group success. Table 7.10 shows the results of probit regressions in which the dependent variable indicates whether the respective group reached the threshold or not. The independent variables *Gender_group*, *Age_group*, *NEP-Index_group*, *Risk_group*, and *Experiments_group* each indicate mean values per group. Model 1 refers to HET-treatments and model 2 to HOM-treatments. We find no clear evidence that any of the independent variables has an effect on group success. The variables *Gender_group* (model 2) and *Risk_group* (model 1) each have a significant negative effect on group success, but only in one model. The variable *Age_group* also has a significant positive effect on group success , but only in model 2.

[49] These are the NEP items 2, 3, 4, 5, 7, 8, 10, 12, 13, 14, and 15.

	(1) HET	(2) HOM
Gender (Female=1)	0.45	0.12
	(0.39)	(0.38)
Age	-0.04	-0.01
	(0.03)	(0.06)
Experiments	0.00	-0.01
	(0.01)	(0.01)
NEP-Index	-0.01	0.00
	(0.03)	(0.03)
Risk	-0.08	0.05
	(0.09)	(0.10)
Rich	1.69***	
	(0.40)	
Constant	4.51***	4.43**
	(1.35)	(1.98)
N	246	114
R^2	0.13	0.01
Adjusted R^2	0.10	-0.05

Table 7.7: OLS regression of contributions in round 11 on demographic characteristics and personal attitudes.

Notes. Both regression models include controls for treatments. Standard errors clustered by group in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Item	Do you agree or disagree that:	Mean (sd)
NEP1	We are approaching the limit of the number of people the earth can support.	4.13 (1.44)
NEP2 (R)	Humans have the right to modify the natural environment to suit their needs.	3.81 (1.13)
NEP3	When humans interfere with nature it often produces disastrous consequences.	4.64 (1.14)
NEP4 (R)	Human ingenuity will insure that we do NOT make the earth unlivable.	3.49 (1.27)
NEP5	Humans are severely abusing the environment.	5.16 (1.02)
NEP6 (R)	The earth has plenty of natural resources if we just learn how to develop them.	2.51 (1.21)
NEP7	Plants and animals have as much right as humans to exist.	4.89 (1.36)
NEP8 (R)	The balance of nature is strong enough to cope with the impacts of modern industrial nations	4.91 (1.11)
NEP9	Despite our special abilities humans are still subject to the laws of nature.	4.96 (1.13)
NEP10 (R)	The so-called "ecological crisis" facing humankind has been greatly exaggerated.	5.30 (1.09)
NEP11	The earth is like a spaceship with very limited room and resources.	3.33 (1.35)
NEP12 (R)	Humans were meant to rule over the rest of nature.	5.06 (1.18)
NEP13	The balance of nature is very delicate and easily upset.	4.54 (1.16)
NEP14 (R)	Humans will eventually learn enough about how nature works to be able to control it.	3.84 (1.28)
NEP15	If things continue on their present course, we will soon experience a major ecological catastrophe.	5.06 (1.03)
All 15		4.38 (0.57)
Answer: (1) Strongly disagree; (2) Somewhat disagree; (3) Rather disagree; (4) Rather agree; (5) Somewhat agree; (6) Strongly agree		

Table 7.8: NEP scale (Dunlop et al., 2000).

Notes. Means and standard deviations (in parentheses). (R) denotes reverse items.

Question	Mean (sd)
How do you see yourself: are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please tick a box on the scale, where the value 0 means: 'not at all willing to take risks' and the value 10 means: 'very willing to take risks'.	5.11 (1.98)

Table 7.9: Question on risk attitude (Dohmen et al., 2011).

Notes. Mean and standard deviation (in parentheses).

	Probit	
	(1)	(2)
	HET	HOM
Gender_group	-2.49	-5.10***
	(1.53)	(1.55)
Age_group	0.00	0.53**
	(0.07)	(0.21)
NEP-Index_group	-0.17	0.35
	(0.10)	(0.25)
Risk_group	-0.58*	0.82
	(0.30)	(0.84)
Experiments_group	-0.05	0.27
	(0.03)	(0.20)
Constant	15.64***	-31.91***
	(5.19)	(10.27)
N	41	19
Wald Chi^2	28.57	17.58
Pseudo R^2	0.29	0.51

Table 7.10: Probit regressions of group success on demographic characteristics and personal attitudes.

Notes. Both regression models include controls for treatments. Standard errors clustered by group in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

8. Conclusion and outlook

The broader question of this dissertation was to what extent economic experiments can contribute to evaluate and potentially increase the efficiency of institutions and technologies that can provide or maintain public goods. The papers in chapters 2, 4 , 6, and 7 each examine specific public goods. In chapters 2 and 4, the efficiency of public goods is evaluated on the basis of the scientific publication process and employer review platforms (ERPs). The papers in chapters 6 and 7 assess how different institutional mechanisms affect the willingness to cooperate in public goods games in the fields of recycling and climate change. The papers in chapters 3 and 5 also provide results that can be used to derive suggestions on how to better maintain and provide public goods within the fields of privacy and environmental economics. In a nutshell, the individual papers of this dissertation show that the experimental method is well suited to evaluate and potentially increase the efficiency of public goods as well as the institutions and technologies that shape them.

Since each paper has its own conclusion, I do not address the main findings of each paper in detail again in this chapter. Instead, I briefly discuss conclusions about the provision and maintenance of public goods that can be drawn from the results of the papers included in this dissertation. Thereby, I formulate further suggestions for future research which were often only briefly discussed in the individual chapters due to space limitations.

The results presented in the paper "Editorial favoritism in the field of laboratory experimental economics" (chapter 2) indicate that the publication process, at least in some top economics journals, does not result in the publication of only the best scientific manuscripts. Instead, the results suggest that social ties between editors and authors also influence editors' decisions. Based on our study, we point out a possible bias in the publication process to the editors of economics journals and suggest how the publication process could be made fairer and more efficient. Since the paper also evaluates methodological standards, experimental economists can use the results of our study as guidance when developing their own experiments. In addition, we indicate what information of an experiment should be reported to allow for replication. Taken together, our findings can thus help to increase credibility and acceptance towards published experimental studies, and facilitate better scientific progress in the field of experimental economics and other disciplines. Consequently, our results can contribute to improving the provision of the public good science.

Future studies could expand research in this and related areas by using additional data sources. In the paper, we solely refer to published papers and therefore are not able to make claims about authors of the submitted manuscripts. For the *American Economic Review* and the *Journal of the European Economic Association*, we find that the share of US authors as well as social ties between editors and authors both have a negative impact on the number of citations a paper receives in the years following publication. However, we do not have data on how many papers were submitted by US authors and non-US authors or by authors with and without social ties to the editors. Card et al. (2020) use submitted manuscripts and published papers to examine whether female and male authors are treated differently in the publication process by editors and referees. Future studies on editorial favoritism could follow the approach of Card et al. (2020) and consider submitted manuscripts in addition to published papers. Basing the examination on all submitted papers would allow to draw more informed conclusions as to whether the selection of published papers is indicative of editorial favoritism.

In the introduction of this dissertation, it was explained that there are only very few publications on the topic of climate change in top economics journals. Using data of submitted manuscripts, similar studies on editorial favoritism could investigate whether topics such as climate change have so far been given little space in top economics journals due to personal preferences (or the research topics) of the editors or, rather, due to a low number of submitted manuscripts on the respective topic. The question here would not only be whether social ties between editors and authors play a role in the selection of submitted manuscripts, but also whether research topics on which an editor herself has conducted research are preferred by the editor in her editorial decisions.

The paper "Is your privacy for sale? An experiment on the willingness to reveal sensitive information" (chapter 3) does not investigate a public good but belongs to the field of basic research. Nevertheless, the results of the paper can be used in other research contexts, e.g. in dealing with the (lack of) provision of public goods. A central result of the paper is that in the case of personal data disclosure, the concrete content of these data has an impact on the willingness to disclose. As an example, in research on recycling of electronic devices, such as smartphones and tablets, this finding can be interpreted as an indication that concerns about disclosure of sensitive data could be a reason for relatively low recycling rates. The results of this paper can also be used to formulate hypotheses about privacy-related decisions on employer review platforms (ERPs) and in smartphone apps, as has been demonstrated in chapters 4 and 5.

The experimental setting presented in the paper could be adapted in future studies on environmental questions. In a first step, information on environmentally harmful and climate-damaging behaviors, such as the number of flights or the use of disposable coffee cups, could be requested. In a second step, the willingness to disclose this information to other participants could be determined. The resulting findings could then be used to formulate recommendations for the promotion of climate-friendly technologies and behaviors.

The results presented in the paper "Employer review platforms - do the rating environment and platform design affect the informativeness of reviews? Theory, evidence, and suggestions" (chapter 4) show that there are a number of factors on ERPs that can affect the informativeness of reviews. Because information is freely available on most ERPs, these platforms can be characterized as a public good in the widest sense. ERPs can reduce information asymmetries between employees and employers. As a result, during their job-finding process, employees may be more likely to identify employers that are suitable for them. Furthermore, in the long run, employees may be more satisfied with their jobs resulting from the job-finding process than they would have been without the information provided on ERPs. In the paper, I identify a number of biasing factors based on the results of economic and psychological experiments and empirically test the effects of these factors on reviews using data from *Kununu* and *Glassdoor*. The effects of biasing factors on reviews are predominantly confirmed by the empirical results which shows that the results of economic experiments are well suited to formulate predictions regarding real life contexts.

The extensive amount of data aggregated on ERPs offers enormous research potential which has so far only been exploited to a relatively small extent. Future studies could, e.g., use the reviews on ERPs to investigate whether evidence of different types of (group-specific) discrimination can be found for specific companies and industries. Furthermore, reviews on ERPs could also be used to develop indices of employee satisfaction. Similar to the ifo Business Climate Index[50], such indices could then be regularly collected and discussed in the media in order to inform the interested public about positive and negative trends regarding employee satisfaction. ERPs could facilitate such research by making the data they collect more easily available. In this way, the societal benefits of ERPs could be increased, potentially sparking the intrinsic motivation of additional employees to submit reviews on ERPs.

[50] See https://www.ifo.de/en/survey/ifo-business-climate-index, accessed April 03, 2021.

The paper "Acceptance of data sharing in smartphone apps from key industries of the digital transformation: A representative population survey for Germany" (chapter 5) provides a method for private and public actors to test the acceptance towards existing and planned technologies. Therefore, the method can help to avoid or correct planning errors and thereby save resources. In treatment 2, the hypothetical smartphone apps are explicitly characterized by the fact that their use can result in positive externalities and thus contribute to the protection of the environment. In the paper, we show to what extent privacy preferences differ according to demographic characteristics and personal attitudes. Our results therefore also provide information that may be of interest for consumer advocates and regulatory agencies.

In the field of environmental economics, future studies could use the method to examine social acceptance towards new climate-friendly technologies and institutional mechanisms. In this way, it would be possible to identify which groups of people are skeptical towards these technologies and institutional mechanisms. In subsequent steps, it could then be investigated which interventions could help to reduce the skepticism of these groups of people.

The results of the paper "Recycling behavior of private households: An empirical investigation of individual preferences in a club good experiment" (chapter 6) allow to cautiously conclude that the provision of the public good recycling can be increased by the introduction of club systems. Recycling activities are at the intersection of local and global public goods. Recycling policies predominantly apply to individual municipalities or countries. However, the benefits of recycling are positive at both the local and global scale. People benefit locally from recycling when, e.g., recycling reduces the amount of waste littering their neighborhoods, or when waste is not burned at the municipal landfill. Globally, people benefit from the fact that the processing and reuse of waste results in a lower resource consumption, which is often associated with lower CO_2 emissions and lower or less rapidly rising resource prices in the long run. The paper's findings may help to develop improvements for local recycling systems that will also have a positive effect at the global level. As explained in the paper, the experiment is a pilot study which will be extended in future studies with more participants and further treatments.

In the paper "Combating climate change: Is the option to exploit a public good a barrier for reaching critical thresholds? Experimental evidence" (chapter 7), we examine an aspect of climate change that has been overlooked by previous experimental literature on the collective risk social dilemma (CRSD). Our results are, in most cases, not significant due to the small number of groups. However, we can observe a trend suggesting that cooperation in the provision of public goods is harder in give-take treatments than in give-only treatments. The

effect of different institutions as well as the effect of uncertain thresholds on the willingness to cooperate in CRSDs has so far only been studied in pure give variants. Therefore, it is not possible to conclude whether previous results also hold true for give-take variants. Thus, there are many additional research opportunities for future studies.

The recycling experiment (chapter 6) and the climate change experiment (chapter 7) were both conducted as part of the third-party funded project "Innovative acceptance research for sustainable development through gamification (GAME)". One goal of the project is to develop a serious game that can be used to investigate decision-making in the fields of environmental and resource economics. In the past, digital games have only been used in few economic studies to generate research data. In these studies, economic games, formerly known from traditional laboratory or field experiments, are embedded in existing virtual worlds, such as *Second Life*, where players act as avatars (Innocenti, 2017; Greiner et al., 2014; Duffy, 2011; Atlas, 2008). However, so far, the economic literature has described few digital serious games in which gamification techniques are explicitly used to generate research data (see e.g., Müller et al. 2017). Due to the global expansion of the internet and the corresponding mobile end devices, well-designed serious games have the potential to generate extensive data material that can be used to test or further develop economic theories.

Unlike in the laboratory experiments in chapters 6 and 7, the game environment in serious games is less artificial and uses so-called immersion. Immersion describes the intensity with which a player feels physically present in the non-physical virtual game world (Blascovich et al., 2002). Through gamification, experimental content can be designed in a more realistic way than in economic laboratory experiments. In addition, serious games can potentially reach larger and more diverse participant pools than traditional laboratory experiments. Since the experiments in chapters 6 and 7 explore public goods whose efficient provision benefits everyone, it can be assumed that the intrinsic motivation of many potential participants is stimulated by embedding such experiments in a serious game and emphasizing the benefits for environmental research. This assumption is also supported by the fact that more than 4.3 million people actively participated in the serious game *Sea Hero Quest*, a mobile game used to collect data for dementia research.[51]

[51] See (in German) https://www.telekom.com/de/verantwortung/details/sea-hero-quest-spielen-gegen-das-vergessen-587114, accessed April 03, 2021.

For experimental economics, an additional benefit of serious games, provided e.g. via a smartphone app, would be that observations could be collected over longer periods of time. In a serious game, the individual rounds from the experiments in chapters 6 and 7 could be played over several days which would give participants the opportunity to think more carefully about their decisions. As described in Czibor et al. (2019), long-term experiments are often avoided by experimental economists because they are associated with higher participation costs and implicit opportunity costs due to the time needed to conduct such experiments. In addition to embedding experiments, a serious game that extends over a longer period would also allow applying the method described in chapter 5 to study acceptance towards climate-friendly technologies and institutional mechanisms.

However, embedding economic experiments in serious games also raises numerous methodological questions, e.g., with regard to the incentivization of the participants and the lower control of the participants compared to the laboratory environment. In order to be able to use the potentials of serious games for economic research in a purposeful way, a critical examination of methodological issues is essential.

References

Abbott, A., Nandeibam, S., & O'Shea, L. (2017). The displacement effect of convenience: the case of recycling. *Ecological Economics*, *136*, 159-168.

Abbott, A., Nandeibam, S., & O'Shea, L. (2013). Recycling: Social norms and warm-glow revisited. *Ecological Economics*, *90*, 10-18.

Abel, J., Klohs, K., Lehmann, H., & Lantow, B. (2017). Sentiment-Analysis for German Employer Reviews. In: *International Conference on Business Information Systems* (pp. 37-48). Springer, Cham.

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2020). Secrets and Likes: The Drive for Privacy and the Difficulty of Achieving It in the Digital Age. *Journal of Consumer Psychology*, *30*(4), 736-758.

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science*, *347*(6221), 509-514.

Acquisti, A., & Gross, R. (2006). Imagined communities: Awareness, information sharing, and privacy on the Facebook. In: *International workshop on privacy enhancing technologies* (pp. 36-58). Springer, Berlin, Heidelberg.

Acquisti, A., & Grossklags, J. (2005). Privacy and rationality in individual decision making. *IEEE security & privacy*, *3*(1), 26-33.

Acquisti, A., Taylor, C., & Wagman, L. (2016). The economics of privacy. *Journal of Economic Literature*, *54*(2), 442-92.

Agarwal, R., Gao, G., DesRoches, C., & Jha, A. K. (2010). Research commentary - The digital transformation of healthcare: Current status and the road ahead. *Information Systems Research*, *21*(4), 796-809.

Akerlof, G. A., & Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, *115*(3), 715-753.

Algan, Y., Benkler, Y., Fuster Morell, M., & Hergueux, J. (2013). Cooperation in a Peer Production Economy Experimental Evidence from Wikipedia. *Available at SSRN 2843518*.

Altmann, S., Milsom, L., Zillessen, H., Blasone, R., Gerdon, F., Bach, R., Kreuter, F., Nosenzo, D., Toussaert, S. & Abeler, J. (2020). Acceptability of app-based contact tracing for COVID-19: Cross-country survey evidence. *Available at SSRN 3590505*.

Amann, J., Sleigh, J., & Vayena, E. (2020). Digital contact-tracing during the Covid-19 pandemic: an analysis of newspaper coverage in Germany, Austria, and Switzerland. *medRxiv*.

Amer, M., Daim, T. U., & Jetter, A. (2013). A review of scenario planning. *Futures*, *46*, 23-40.

Andone, I., Błaszkiewicz, K., Eibes, M., Trendafilov, B., Montag, C., & Markowetz, A. (2016). How age and gender affect smartphone usage. In: *Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing: adjunct* (pp. 9-12). ACM.

Andreoni, J. (1995). Warm-glow versus cold-prickle: the effects of positive and negative framing on cooperation in experiments. *The Quarterly Journal of Economics*, *110*(1), 1-21.

Andreoni, J. (1988). Why free ride?: Strategies and learning in public goods experiments. *Journal of Public Economics*, *37*(3), 291-304.

Angus, S. D., Atalay, K., Newton, J., & Ubilava, D. (2020). Geographic Diversity in Economic Publishing. *Available at SSRN 3697906*.

Antonioni, D. (1994). The effects of feedback accountability on upward appraisal ratings. *Personnel Psychology*, *47*(2), 349-356.

App, S., Merk, J., & Büttgen, M. (2012). Employer branding: Sustainable HRM as a competitive advantage in the market for high-quality employees. *mrev management revue, 23*(3), 262-278.

Apthorpe, N., Shvartzshnaider, Y., Mathur, A., Reisman, D., & Feamster, N. (2018). Discovering smart home internet of things privacy norms using contextual integrity. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *2*(2), 1-23.

Ariely, D., Bracha, A., & Meier, S. (2009). Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *American Economic Review*, *99*(1), 544-555.

Arnold, A., & Staffelbach, B. (2012). Perceived post-restructuring job insecurity: The impact of employees' trust in one's employer and perceived employability. *German Journal of Human Resource Management*, *26*(4), 307-330.

Asensio, O. I., & Delmas, M. A. (2016). The dynamics of behavior change: Evidence from energy conservation. *Journal of Economic Behavior & Organization*, *126*, 196-212.

Asensio, O. I., & Delmas, M. A. (2015). Nonprice incentives and energy conservation. *Proceedings of the National Academy of Sciences*, *112*(6), 510-515.

Atlas, S. (2008). Inductive metanomics: economic experiments in virtual worlds. *Journal of Virtual Worlds Research*, *1*(1), 1–15.

Ayres, I., Banaji, M., & Jolls, C. (2015). Race effects on eBay. *The RAND Journal of Economics*, *46*(4), 891-917.

Bardsley, N. (2008). Dictator game giving: altruism or artefact?. *Experimental Economics*, *11*(2), 122-133.

Bardsley, N., Cubitt, R., Loomes, G., Moffatt, P., Starmer, C., & Sugden R. (2010). *Experimental Economics - Rethinking the Rules.* Princeton University Press.

Barnes, N. G., Lescault, A. M., & Holmes, G. (2015). The 2015 Fortune 500 and social media: Instagram gains, blogs lose. *University of Massachusetts at Dartmouth–Center for Marketing Research*.

Barr, S. (2007). Factors influencing environmental attitudes and behaviors: A UK case study of household waste management. *Environment and Behavior*, *39*(4), 435-473.

Barrett, S., & Dannenberg, A. (2014). Sensitivity of collective action to uncertainty about climate tipping points. *Nature Climate Change*, *4*(1), 36-39.

Barrett, S., & Dannenberg, A. (2012). Climate negotiations under scientific uncertainty. *Proceedings of the National Academy of Sciences*, *109*(43), 17372-17376.

Becker, G. M., DeGroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, *9*(3), 226-232.

Beede, D. N., & Bloom, D. E. (1995). The economics of municipal solid waste. *The World Bank Research Observer*, *10*(2), 113-150.

Bellemare, C., Bissonnette, L., & Kröger S. (2016). Simulating power of economic experiments: the powerBBK package. *Journal of the Economic Science Association*, 2(2), 157-168.

Bellemare, C., Bissonnette, L., & Kröger S. (2014). Statistical power of within and between-subjects designs in economic experiments. IZA Discussion Paper No. 8583.

Benkler, Y. (2006). *The wealth of networks: How social production transforms markets and freedom*. Yale University Press.

Benndorf, V. (2018). Voluntary disclosure of private information and unraveling in the market for lemons: An experiment. *Games*, *9*(2), 23.

Benndorf, V., & Normann, H. T. (2018). The willingness to sell personal data. *The Scandinavian Journal of Economics*, *120*(4), 1260-1278.

Beresford, A. R., Kübler, D., & Preibusch, S. (2012). Unwillingness to pay for privacy: A field experiment. *Economics Letters*, *117*(1), 25-27.

Berger, J. (2011). Arousal increases social transmission of information. *Psychological Science*, *22*(7), 891-893.

Berger, J., & Milkman, K.L. (2012). What makes online content viral?. *Journal of Marketing Research*, *49*(2), 192-205.

Bidwell, M., Won, S., Barbulescu, R., & Mollick, E. (2015). I used to work at Goldman Sachs! How firms benefit from organizational status in the market for human capital. *Strategic Management Journal*, *36*(8), 1164-1173.

Bitkom (2020). Bitkom veranstaltet erstmals die Digital Transformation Week. Online: https://www.bitkom.org/Presse/Presseinformation/Bitkom-veranstaltet-erstmals-die-Digital-Transformation-Week (accessed October 08, 2020).

Bitzer, J., Schrettl, W., & Schröder, P.J. (2007). Intrinsic motivation in open source software development. *Journal of Comparative Economics*, *35*(1), 160-169.

Blackwell, C., & McKee, M. (2003). Only for my own neighborhood?: Preferences and voluntary provision of local and global public goods. *Journal of Economic Behavior & Organization*, *52*(1), 115-131.

Blank, G., Bolsover, G., & Dubois, E. (2014). A new privacy paradox: Young people and privacy on social network sites. In: *Proceedings of the Annual Meeting of the American Sociological Association, San Francisco, CA, USA, 17 August 2014*.

Blascovich, J., Loomis, J., Beall, A. C., Swinth, K. R., Hoyt, C. L., & Bailenson, J. N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry*, *13*(2), 103-124.

Blondeel, M., & Van de Graaf, T. (2018). Toward a global coal mining moratorium? A comparative analysis of coal mining policies in the USA, China, India and Australia. *Climatic Change*, *150*(1), 89-101.

Bock, O., Baetge, I., & Nicklisch, A. (2014). hroot: Hamburg registration and organization online tool. *European Economic Review*, *71*, 117-120.

Bohm, P., Lindén, J., & Sonnegård, J. (1997). Eliciting reservation prices: Becker–DeGroot–Marschak mechanisms vs. markets. *The Economic Journal*, *107*(443), 1079-1089.

Böhm, R., Gürerk, Ö., & Lauer, T. (2020). Nudging Climate Change Mitigation: A Laboratory Experiment with Inter-Generational Public Goods. *Games*, *11*(4), 42.

Bolton, G., Greiner, B., & Ockenfels, A. (2013). Engineering trust: reciprocity in the production of reputation information. *Management Science*, *59*(2), 265-285.

Brekke, K. A., Kipperberg, G., & Nyborg, K. (2010). Reluctant recyclers: social interaction in responsibility ascription. *Land Economics, 86*, 766-784.

Briguglio, M. (2016). Household cooperation in waste management: initial conditions and intervention. *Journal of Economic Surveys*, *30*(3), 497-525.

Brogaard, J., Engelberg, J., & Parsons, C. A. (2014). Networks and productivity: Causal evidence from editor rotations. *Journal of Financial Economics*, 111(1), 251-270.

Brown, T. C., & Kroll, S. (2017). Avoiding an uncertain catastrophe: climate change mitigation under risk and wealth heterogeneity. *Climatic Change*, *141*(2), 155-166.

Brown-Kruse, J., & Hummels, D. (1993). Gender effects in laboratory public goods contribution: Do individuals put their money where their mouth is?. *Journal of Economic Behavior & Organization*, *22*(3), 255-267.

Brutus, S., & Derayeh, M. (2002). Multisource assessment programs in organizations: An insider's perspective. *Human Resource Development Quarterly*, *13*(2), 187-202.

Bucher, D., Mangili, F., Cellina, F., Bonesana, C., Jonietz, D., & Raubal, M. (2019). From location tracking to personalized eco-feedback: A framework for geographic information collection, processing and visualization to promote sustainable mobility behaviors. *Travel Behaviour and Society*, *14*, 43-56.

Bundesministerium für Gesundheit (2019). Daten des Gesundheitswesens. Online: https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/5_Publikationen/Gesundheit/Broschueren/BMG_DdGW_2019_bf.pdf (accessed December 01, 2020).

Bundesministerium für Wirtschaft und Energie (2020). Smart Meter und digitale Stromzähler. Eine sichere digitale Infrastruktur für die Energiewende. Online: https://www.bmwi.de/Redaktion/DE/Publikationen/Energie/smart-meter-und-digitale-stromzaehler.pdf?__blob=publicationFile&v=12 (accessed October 13, 2020).

Burton-Chellew, M. N., May, R. M., & West, S. A. (2013). Combined inequality in wealth and risk leads to disaster in the climate change game. *Climatic Change*, *120*(4), 815-830.

Camerer, C. F., Dreber, A., Forsell, E., Ho, T. H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T., Razen, M., & Wu, H. (2016), Evaluating replicability of laboratory experiments in economics. *Science*, *351*(6280), 1433-1436.

Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T. H., Huber, J., Johannesson, M., Kirchler, M., Nave, G., Nosek, B. A., Pfeiffer, T., Altmejd, A., Buttrick, N., Chan, T., Chen, Y., Forsell, E., Gampa, A., Heikensten, E., Hummer, L., Imai, T., Isaksson, S., Manfredi, D., Rose, J., Wagenmakers, E.-J., & Wu, H. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour*, *2*(9), 637-644.

Camerer, C. F., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*, 19(1-3), 7-42.

Card, D., & DellaVigna, S. (2020). What do editors maximize? Evidence from four economics journals. *Review of Economics and Statistics*, 102(1), 195-217.

Card, D., DellaVigna, S., Funk, P., & Iriberri, N. (2020). Are referees and editors in economics gender neutral?. *The Quarterly Journal of Economics*, 135(1), 269-327.

Cecere, G., Mancinelli, S., & Mazzanti, M. (2014). Waste prevention and social preferences: the role of intrinsic and extrinsic motivations. *Ecological Economics*, *107*, 163-176.

Cellina, F., Bucher, D., Mangili, F., Veiga Simão, J., Rudel, R., & Raubal, M. (2019). A large scale, app-based behaviour change experiment persuading sustainable mobility patterns: Methods, results and lessons learnt. *Sustainability*, *11*(9), 2674.

Chakravarty, S., & Fonseca, M. A. (2017). Discrimination via exclusion: An experiment on group identity and club goods. *Journal of Public Economic Theory*, *19*(1), 244-263.

Chapman, D.S., Uggerslev, K.L., Carroll, S.A., Piasentin, K.A., & Jones, D.A. (2005). Applicant attraction to organizations and job choice: a meta-analytic review of the correlates of recruiting outcomes. *Journal of Applied Psychology*, *90*(5), 928-944.

Chatelan, A., Bochud, M., & Frohlich, K. L. (2019). Precision nutrition: hype or hope for public health interventions to reduce obesity?. *International Journal of Epidemiology*, *48*(2), 332-342.

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics*, *14*(1), 47-83.

Chavanne, D. (2018). Generalized trust, need for cognitive closure, and the perceived acceptability of personal data collection. *Games*, *9*(2), 18.

Chen, Y., Li, S. X., Liu, T. X., & Shih, M. (2014). Which hat to wear? Impact of natural identities on coordination and cooperation. *Games and Economic Behavior*, *84*, 58-86.

Chen, Y., & Li, S. X. (2009). Group identity and social preferences. *American Economic Review*, *99*(1), 431-57.

Cherry, T. L., & Dickinson, D. L. (2008). Voluntary contributions with multiple public goods. In: Cherry, T. L., Kroll, S., & Shogren, J. F., *Environmental economics, experimental methods*, (pp. 184-193), Routledge.

Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, *55*, 591-621.

Cloos, J. (2021). Employer Review Platforms–Do the Rating Environment and Platform Design affect the Informativeness of Reviews? Theory, Evidence, and Suggestions. *mrev management revue*, *32*(3), 152-181.

Cloos, J., Frank, B., Kampenhuber, L., Karam, S., Luong, N., Möller, D., Monge-Larrain, M., Tan Dat, N., Nilgen, M., & Rössler, C. (2019). Is Your Privacy for Sale? An Experiment on the Willingness to Reveal Sensitive Information. *Games*, *10*(3), 28.

Cloos, J., Greiff, M., & Rusch, H. (2020). *Geographical Concentration and Editorial Favoritism within the Field of Laboratory Experimental Economics (RM/19/029-revised-)*. Maastricht University, Graduate School of Business and Economics. GSBE Research Memoranda, No. 014.

Collier, P. (2019). Sir Paul Collier Recommends “Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being” by George A. Akerlof and Rachel E. Kranton. In: *21st century economics* (pp. 29-31). Springer, Cham.

Colussi, T. (2018). Social ties in academia: A friend is a treasure. *Review of Economics and Statistics*, 100(1), 45-50.

Conte, F. (2018). Understanding the influence of CEO tenure and CEO reputation on corporate reputation: An exploratory study in Italy. *International Journal of Business and Management*, *13*(3), 54-66.

Cornes, R., & Sandler, T. (1996). *The theory of externalities, public goods, and club goods*. Cambridge University Press.

Cortiñas, M., Elorz, M., & Múgica, J. M. (2008). The use of loyalty-cards databases: Differences in regular price and discount sensitivity in the brand choice decision between card and non-card holders. *Journal of Retailing and Consumer Services*, *15*(1), 52-62.

Craciun, G., & Moore, K. (2019). Credibility of negative online product reviews: Reviewer gender, reputation and emotion effects. *Computers in Human Behavior*, *97*, 104-115.

Cui, R., Li, J., & Zhang, D. J. (2020). Reducing discrimination with reviews in the sharing economy: Evidence from field experiments on Airbnb. *Management Science*, *66*(3), 1071-1094.

Czibor, E., Jimenez-Gomez, D., & List, J. A. (2019). The dozen things experimental economists should do (more of). *Southern Economic Journal*, *86*(2), 371-432.

Dabirian, A., Kietzmann, J., & Diba, H. (2017). A great place to work!? Understanding crowdsourced employer branding. *Business Horizons*, *60*(2), 197-205.

Dannenberg, A., & Gallier, C. (2019). The choice of institutions to solve cooperation problems: a survey of experimental research. *Experimental Economics*, 1-34.

Dannenberg, A., Löschel, A., Paolacci, G., Reif, C., & Tavoni, A. (2015). On the provision of public goods with probabilistic and ambiguous thresholds. *Environmental and Resource Economics*, *61*(3), 365-383.

Dannenberg, A., & Tavoni, A. (eds.). (2016). Collective Action in Dangerous Climate Change Games. In: *WORLD SCIENTIFIC REFERENCE ON NATURAL RESOURCES AND ENVIRONMENTAL POLICY IN THE ERA OF GLOBAL CHANGE: Volume 4: Experimental Economics* (pp. 95-120).

Davis, D. D., & Holt, C. A. (1993). *Experimental economics*. Princeton University Press.

Dawid, H., & Gatti, D. D. (2018). Agent-based macroeconomics. *Handbook of Computational Economics*, *4*, 63-156.

Del Vecchio, P., Secundo, G., Maruccia, Y., & Passiante, G. (2019). A system dynamic approach for the smart mobility of people: Implications in the age of big data. *Technological Forecasting and Social Change*, *149*, 119771.

DiFilippo, K. N., Huang, W. H., Andrade, J. E., & Chapman-Novakofski, K. M. (2015). The use of mobile apps to improve nutrition outcomes: a systematic literature review. *Journal of Telemedicine and Telecare*, *21*(5), 243-253.

Do Paço, A. M. F., & Reis, R. (2012). Factors affecting skepticism toward green advertising. *Journal of Advertising*, *41*(4), 147-155.

Dogruel, L., Joeckel, S., & Vitak, J. (2017). The valuation of privacy premium features for smartphone apps: The influence of defaults and expert recommendations. *Computers in Human Behavior*, *77*, 230-239.

Dohmen, T., Falk, A., Golsteyn, B. H., Huffman, D., & Sunde, U. (2017). Risk attitudes across the life course. *The Economic Journal, 127*(605), 95-116.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, *9*(3), 522-550.

Doleac, J. L., & Stein, L. C. (2013). The visible hand: Race and online market outcomes. *The Economic Journal*, *123*(572), 469-492.

Dolnicar, S., & Jordaan, Y. (2007). A market-oriented approach to responsibly managing information privacy concerns in direct marketing. *Journal of Advertising*, *36*(2), 123-149.

Dorner, V., Giamattei, M., & Greiff, M. (2020). The Market for Reviews: Strategic Behavior of Online Product Reviewers with Monetary Incentives. *Schmalenbach Business Review*, *72*(3), 397-435.

Drabe, D., Hauff, S., & Richter, N.F. (2015). Job satisfaction in aging workforces: an analysis of the USA, Japan and Germany. *The International Journal of Human Resource Management*, *26*(6), 783-805.

Duffy, J. (2011). Trust in second life. *Southern Economic Journal*, *78*(1), 53-62.

Duflo, E. (2020). Report of the Editor: American Economic Review. *AEA Papers and Proceedings*, *110*, 660-74.

Dufwenberg, M., Gächter, S., & Hennig-Schmidt, H. (2011). The framing of games and the psychology of play. *Games and Economic Behavior*, *73*(2), 459-478.

Dunlap, R. E., Van Liere, K. D., Mertig, A. G., & Jones, R. E. (2000). New trends in measuring environmental attitudes: measuring endorsement of the new ecological paradigm: a revised NEP scale. *Journal of Social Issues*, *56*(3), 425-442.

Edelman, B., Luca, M., & Svirsky, D. (2017). Racial discrimination in the sharing economy: Evidence from a field experiment. *American Economic Journal: Applied Economics*, *9*(2), 1-22.

Egelman, S., Felt, A. P., & Wagner, D. (2013). Choice architecture and smartphone privacy: There'sa price for that. In: *The Economics of Information Security and Privacy* (pp. 211-236). Springer, Berlin, Heidelberg.

Einav, L., & Levin, J. (2014). Economics in the age of big data. *Science*, *346*(6210).

Erlei, M., Leschke, M., & Sauerland, D. (2016). *Neue Institutionenökonomik*. Schäffer-Poeschel.

Falk, A., & Fehr, E. (2003). Why labour market experiments?. *Labour Economics*, *10*(4), 399-406.

Falk, A., & Heckman, J. J. (2009). Lab experiments are a major source of knowledge in the social sciences. *Science*, *326*(5952), 535-538.

Fast, V., & Schnurr, D. (2020). The Value of Personal Data: An Experimental Analysis of Data Types and Personal Antecedents. *Available at SSRN 3611232*.

Federal Statistical Office of Germany (2019). Labour market, Employment rate. Retrieved from: https://www.destatis.de/EN/Themes/Economy/Short-Term-Indicators/_node.html (accessed December 29, 2019).

Fehr, E., & Gächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives*, *14*(3), 159-181.

Fehr, E., & Gächter, S. (1998). Reciprocity and economics: The economic implications of homo reciprocans. *European Economic Review*, *42*(3-5), 845-859.

Fernández-Manzanal, R., Rodríguez-Barreiro, L., & Carrasquer, J. (2007). Evaluation of environmental attitudes: Analysis and results of a scale applied to university students. *Science Education*, *91*(6), 988-1009.

Fernández-Rovira, C., Valdés, J. Á., Molleví, G., & Nicolas-Sans, R. (2021). The digital transformation of business. Towards the datafication of the relationship with customers. *Technological Forecasting and Social Change*, *162*, 120339.

Fietze, S., Matiaske, W., & Menges, R. (2019). Corporate responsibility: In the dilemma between fake and trust?. *mrev management revue*, *30*(2-3), 143-147.

Filippas, A., Horton, J.J., & Golden, J. (2018). Reputation inflation. In *Proceedings of the 2018 ACM Conference on Economics and Computation* (pp. 483-484). ACM.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171-178.

Fogel, J., & Nehmad, E. (2009). Internet social network communities: Risk taking, trust, and privacy concerns. *Computers in Human Behavior*, *25*(1), 153-160.

Foss, N. J. (2019). Nicolai J. Foss Recommends "Economics and Identity" by George A. Akerlof and Rachel E. Kranton. In: *21st Century Economics* (pp. 49-50). Springer, Cham.

Fréchette, G. R., & Schotter, A. (eds.). (2015). *Handbook of experimental economic methodology*. The Handbooks of Economic Methodol Series, New York University.

Frederick, S., & Loewenstein, G. (1999). 16 Hedonic Adaptation. In: Kahneman, D., Diener, E., & Schwarz, N. (eds.), *Well-Being: The Foundations of Hedonic Psychology* (pp. 302-329). Rusell Sage Foundation.

Friedman, S., Friedman, D., & Sunder, S. (1994). *Experimental methods: A primer for economists*. Cambridge University Press.

Fuller, C. S. (2019). Is the market for digital privacy a failure?. *Public Choice*, *180*(3), 353-381.

Gächter, S., Kölle, F., & Quercia, S. (2017). Reciprocity and the tragedies of maintaining and providing the commons. *Nature Human Behaviour*, *1*(9), 650-656.

Gadenne, V. (2011). Über die Validität von Experimenten in der Ökonomie. In: Gadenne, V., & Neck, R. (eds.), *Philosophie und Wirtschaftswissenschaft* (pp. 51-64). Tübingen: Mohr Siebeck.

Ghosh, A., & Roth, A. (2015). Selling privacy at auction. *Games and Economic Behavior*, *91*, 334-346.

Giamattei, M., & Lambsdorff, J. G. (2019). classEx—an online tool for lab-in-the-field experiments with smartphones. *Journal of Behavioral and Experimental Finance*, *22*, 223-231.

Gisdakis, S., Manolopoulos, V., Tao, S., Rusu, A., & Papadimitratos, P. (2014). Secure and privacy-preserving smartphone-based traffic information systems. *IEEE Transactions on intelligent transportation systems*, *16*(3), 1428-1438.

Goldfarb, A., & Tucker, C. (2012). Shifts in privacy concerns. *American Economic Review*, *102*(3), 349-53.

Gollier, C. & Tirole, J. (2017). Negotiating effective institutions against climate change. In: Cramton, P., MacKay, D. J., Ockenfels, A. & Stoft, S. (eds.), *Global Climate Pricing: The Path to Climate Protection* (pp. 165-204). Cambridge, MA: MIT Press.

Greiff, M., & Cloos, J. (2019). Experimentelle Entdeckungsreisen. *WiSt-Wirtschaftswissenschaftliches Studium*, *48*(7-8), 32-37.

Greiff, M., & Paetzel, F. (2020). Information about average evaluations spurs cooperation: An experiment on noisy reputation systems. *Journal of Economic Behavior & Organization*, *180*, 334-356.

Greiner, B., Caravella, M., & Roth, A. E. (2014). Is avatar-to-avatar communication as effective as face-to-face communication? An Ultimatum Game experiment in First and Second Life. *Journal of Economic Behavior & Organization*, *108*, 374-382.

Greveler, U., Glösekötterz, P., Justusy, B., & Loehr, D. (2012). Multimedia content identification through smart meter power usage profiles. In: *Proceedings of the International Conference on Information and Knowledge Engineering (IKE)* (p. 1).

Grund, C. (2006). Mitarbeiterrekrutierung über das Internet - Marktanalyse und empirische Untersuchung von Determinanten und Konsequenzen für die Arbeitnehmer. *Journal of Business Economics*, *76*(5), 451-472.

Guido, G., Vitale, A., Astarita, V., Saccomanno, F., Giofré, V. P., & Gallelli, V. (2012). Estimation of safety performance measures from smartphone sensors. *Procedia-Social and Behavioral Sciences*, *54*, 1095-1103.

Gürerk, Ö., Irlenbusch, B., & Rockenbach, B. (2006). The competitive advantage of sanctioning institutions. *Science*, *312*(5770), 108-111.

Ham, C. D. (2017). Exploring how consumers cope with online behavioral advertising. *International Journal of Advertising*, *36*(4), 632-658.

Hamermesh, D. S. (2018). Citations in economics: Measurement, uses, and impacts. *Journal of Economic Literature*, 56(1), 115-56.

Hamermesh, D. S. (1994). Facts and myths about refereeing. *Journal of Economic Perspectives*, 8(1), 153-163.

Hann, I. H., Hui, K. L., Lee, S. Y. T., & Png, I. P. (2007). Overcoming online information privacy concerns: An information-processing theory approach. *Journal of Management Information Systems*, *24*(2), 13-42.

Harrison, G. W., & List, J. A. (2004). Field experiments. *Journal of Economic Literature*, *42*(4), 1009-1055.

Haws, K. L., Winterich, K. P., & Naylor, R. W. (2014). Seeing the world through GREEN-tinted glasses: Green consumption values and responses to environmentally friendly products. *Journal of Consumer Psychology*, *24*(3), 336-354.

Heider, F. (1946). Attitudes and cognitive organization. *The Journal of Psychology*, *21*(1), 107-112.

Helm, S. (2011). Employees' awareness of their impact on corporate reputation. *Journal of Business Research*, *64*(7), 657-663.

Hengel, E., & Moon, E. (2020). *Gender and quality at top economics journals*. Working Paper, University of Liverpool, UK.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010a). The weirdest people in the world?. *Behavioral and Brain Sciences*, *33*(2-3), 61-83.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010b). Most people are not WEIRD. *Nature*, *466*(7302), 29-29.

Henry, L. A., & Möllering, G. (2019). Collective Corporate Social Responsibility: The Role of Trust as an Organizing Principle. *mrev management revue*, *30*(2-3), 173-191.

Hertwig, R., & Ortmann, A. (2001). Experimental practices in economics: A methodological challenge for psychologists?. *Behavioral and Brain Sciences*, *24*(3), 383-403.

Hinz, O., Nofer, M., Schiereck, D., & Trillig, J. (2015). The influence of data theft on the share prices and systematic risk of consumer electronics companies. *Information & Management*, *52*(3), 337-347.

Hoffman, E., McCabe, K. A., & Smith, V. L. (1998). Behavioral foundations of reciprocity: Experimental economics and evolutionary psychology. *Economic Inquiry*, *36*(3), 335-352.

Hoon, C., Bormann, K. C., Graffius, M., & Hansen, C. (2019). The Impact of Organizational Scandals on Employee Voice Behaviors. In: *Academy of Management Proceedings* (Vol. 2019, No. 1, p. 17206). Briarcliff Manor, NY 10510: Academy of Management.

Horne, C., Darras, B., Bean, E., Srivastava, A., & Frickel, S. (2015). Privacy, technology, and norms: The case of smart meters. *Social Science Research*, *51*, 64-76.

Hu, N., Zhang, J., & Pavlou, P.A. (2009). Overcoming the J-shaped distribution of product reviews. *Communications of the ACM*, *52*(10), 144-147.

Huang, Y., Kendrick, K.M., & Yu, R. (2014). Conformity to the opinions of other people lasts for no more than 3 days. *Psychological Science*, *25*(7), 1388-1393.

Huang, K., Li, M., & Markov, S. (2020). What do employees know? Evidence from a social media platform. *The Accounting Review*, *95*(2), 199-226.

Huang, M., Li, P., Meschke, F., & Guthrie, J.P. (2015). Family firms, employee satisfaction, and corporate performance. *Journal of Corporate Finance*, *34*, 108-127.

Huberman, B. A., Adar, E., & Fine, L. R. (2005). Valuating privacy. *IEEE security & privacy*, *3*(5), 22-25.

Huhtala, A. (2010). Income effects and the inconvenience of private provision of public goods for bads: The case of recycling in Finland. *Ecological Economics*, *69*(8), 1675-1681.

Inman, J. J., & Nikolova, H. (2017). Shopper-facing retail technology: A retailer adoption decision framework incorporating shopper attitudes and privacy concerns. *Journal of Retailing*, *93*(1), 7-28.

Innocenti, A. (2017). Virtual reality experiments in economics. *Journal of Behavioral and Experimental Economics*, *69*, 71-77.

Ioannidis, J. P. A., Stanley, T. D., & Doucouliagos, H. (2017). The power of bias in economics research. *The Economic Journal*, 127(605), 236-265.

IPCC (2018). Summary for Policymakers. In: Masson-Delmotte, V., P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J.B.R. Matthews, Y. Chen, X. Zhou, M.I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, and T. Waterfield (eds.), *Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate*. World Meteorological Organization, Geneva, Switzerland, 32 pp.

Iqbal, M. U., & Lim, S. (2010). Privacy implications of automated GPS tracking and profiling. *IEEE Technology and Society Magazine*, *29*(2), 39-46.

Isaac, R. M., & Walker, J. M. (1988). Group size effects in public goods provision: The voluntary contributions mechanism. *The Quarterly Journal of Economics*, *103*(1), 179-199.

Izuma, K., & Adolphs, R. (2013). Social manipulation of preference in the human brain. *Neuron*, *78*(3), 563-573.

Jacquet, J., Hagel, K., Hauert, C., Marotzke, J., Röhl, T., & Milinski, M. (2013). Intra-and intergenerational discounting in the climate game. *Nature Climate Change*, *3*(12), 1025-1028.

Jakob, M., Steckel, J. C., Jotzo, F., Sovacool, B. K., Cornelsen, L., Chandra, R., Edenhofer, O., Holden, C., Löschel, A., Nace, T., Robins, N., Suedekum, J. & Urpelainen, J. (2020). The future of coal in a carbon-constrained climate. *Nature Climate Change*, *10*(8), 704-707.

Jin, H., Su, L., Xiao, H., & Nahrstedt, K. (2018). Incentive mechanism for privacy-aware data aggregation in mobile crowd sensing systems. *IEEE/ACM Transactions on Networking*, *26*(5), 2019-2032.

Jochims, T. (2016). Social reciprocity as a critical success factor for small and mid-size enterprises: Work relationships as reflections of social exchange structures. *mrev management revue*, *27*(3), 188-207.

John, L. K., Acquisti, A., & Loewenstein, G. (2011). Strangers on a plane: Context-dependent willingness to divulge sensitive information. *Journal of Consumer Research*, *37*(5), 858-873.

Johnson, J. P. (2013). Targeted advertising and advertising avoidance. *The RAND Journal of Economics*, *44*(1), 128-144.

Joy, A., Stahl, T., Mohnike, C., Jirage, R., Kula, K. D., & Daim, T. U. (2020). Ethical Issues of Data Tracking and Analytics. In: Daim, T. U. & Meissner, D. (eds.), *Innovation Management in the Intelligent World* (pp. 81-97). Springer, Cham.

Kagel J. H., & Levin D. (2010). Auctions (experiments). In: Durlauf S. N., Blume L. E. (eds), *Behavioural and Experimental Economics* (pp. 14-22). The New Palgrave Economics Collection. Palgrave Macmillan, London.

Karousakis, K., & Birol, E. (2008). Investigating household preferences for kerbside recycling services in London: A choice experiment approach. *Journal of Environmental Management*, *88*(4), 1099-1108.

Khadjavi, M., & Lange, A. (2015). Doing good or doing harm: experimental evidence on giving and taking in public good games. *Experimental Economics*, *18*(3), 432-441.

Kim, J., & Gupta, P. (2012). Emotional expressions in online user reviews: How they influence consumers' product evaluations. *Journal of Business Research*, *65*(7), 985-992.

Kinnaman, T. C. (2009). The economics of municipal solid waste management. *Waste Management*, 2615.

Kokolakis, S. (2017). Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & Security*, *64*, 122-134.

Kollitz, R., Ruhle, S., & Süß, S. (2019). Recruitment practices under scrutiny: A latent-profile analysis of family firms' approaches to recruit non-family employees. *German Journal of Human Resource Management*, *33*(3), 167-196.

Könsgen, R., Schaarschmidt, M., Ivens, S., & Munzel, A. (2018). Finding meaning in contradiction on employee review sites—effects of discrepant online reviews on job application intentions. *Journal of Interactive Marketing*, *43*, 165-177.

Kreps, D. M. (1997). Intrinsic motivation and extrinsic incentives. *American Economic Review*, *87*(2), 359-364.

Krupka, E. L., & Weber, R. A. (2013). Identifying social norms using coordination games: Why does dictator game sharing vary?. *Journal of the European Economic Association*, *11*(3), 495-524.

Laband, D. N., & Piette, M. J. (1994a). Favoritism versus search for good papers: Empirical evidence regarding the behavior of journal editors. *Journal of Political Economy*, *102*(1), 194-203.

Laband, D. N., & Piette, M. J. (1994b). The relative impacts of economics journals: 1970-1990. *Journal of Economic Literature*, *32*(2), 640-666.

Landers, R. N., Brusso, R. C., & Auer, E. M. (2019). Crowdsourcing Job Satisfaction Data: Examining the Construct Validity of Glassdoor. com Ratings. *Personnel Assessment and Decisions*, *5*(3), 45-55.

Ledyard, J. O. (1995). Public goods: a survey of experimental research. In: Kagel J. H., Roth A. E. (eds.), *Handbook of experimental economics* (pp. 111–194). Princeton University Press, Princeton.

Levitt, S. D., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world?. *Journal of Economic Perspectives*, *21*(2), 153-174.

List, J. A. (2007). On the interpretation of giving in dictator games. *Journal of Political Economy*, *115*(3), 482-493.

List, J. A., Sadoff, S., & Wagner, M. (2011). So you want to run an experiment, now what? Some simple rules of thumb for optimal experimental design. *Experimental Economics*, *14*(4), 439-457.

Lohse, J., & Waichman, I. (2020). The effects of contemporaneous peer punishment on cooperation with the future. *Nature Communications*, *11*(1), 1-8.

Longhi, L., & Nanni, M. (2020). Car telematics big data analytics for insurance and innovative mobility services. *Journal of Ambient Intelligence and Humanized Computing*, *11*, 3989-3999.

Love, E. G., Lim, J., & Bednar, M. K. (2017). The face of the firm: The influence of CEOs on corporate reputation. *Academy of Management Journal*, *60*(4), 1462-1481.

Luca, M. (2017), Designing Online Marketplaces: Trust and Reputation Mechanisms. *Innovation Policy and the Economy*, *17*(1), 77-93.

Luca, M., & Zervas, G. (2016). Fake it till you make it: Reputation, competition, and Yelp review fraud. *Management Science*, *62*(12), 3412-3427.

Luo, N., Zhou, Y., & Shon, J. (2016). Employee satisfaction and corporate performance: Mining employee reviews on Glassdoor.com. In: *Thirty Seventh International Conference on Information Systems*, Dublin.

Marinescu, I., Chamberlain, A., Smart, M., & Klein, N. (2021). Incentives can reduce bias in online employer reviews. *Journal of Experimental Psychology: Applied*.

Masterov, D.V., Mayer, U.F., & Tadelis, S. (2015). Canary in the e-commerce coal mine: Detecting and predicting poor experiences using buyer-to-seller messages. In: *Proceedings of the Sixteenth ACM Conference on Economics and Computation* (pp. 81-93). ACM.

Mayzlin, D., Dover, Y., & Chevalier, J. (2014). Promotional reviews: An empirical investigation of online review manipulation. *American Economic Review*, *104*(8), 2421-2455.

Medoff, M. H. (2003). Editorial favoritism in economics?. *Southern Economic Journal*, *70*(2), 425-434.

Mehta, N., & Pandit, A. (2018). Concurrence of big data analytics and healthcare: A systematic review. *International Journal of Medical Informatics*, *114*, 57-65.

Men, L. R. (2012). CEO credibility, perceived organizational reputation, and employee engagement. *Public Relations Review*, *38*(1), 171-173.

Menges, R., Schroeder, C., & Traub, S. (2005). Altruism, warm glow and the willingness-to-donate for green electricity: an artefactual field experiment. *Environmental and Resource Economics*, *31*(4), 431-458.

Milinski, M., Röhl, T., & Marotzke, J. (2011). Cooperative interaction of rich and poor can be catalyzed by intermediate climate targets. *Climatic Change*, *109*(3), 807-814.

Milinski, M., Sommerfeld, R. D., Krambeck, H. J., Reed, F. A., & Marotzke, J. (2008). The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proceedings of the National Academy of Sciences*, *105*(7), 2291-2294.

Minch, R. P. (2015). Location privacy in the era of the internet of things and big data analytics. In: *2015 48th Hawaii International Conference on System Sciences* (pp. 1521-1530). IEEE.

Mitchard, E. T. (2018). The tropical forest carbon cycle and climate change. *Nature*, *559*(7715), 527-534.

Moed, H. F. (2006). *Citation analysis in research evaluation*. Vol. 9. Springer Science & Business Media.

Mögele, B., & Tropp, J. (2010). The emergence of CSR as an advertising topic: a longitudinal study of German CSR advertisements. *Journal of Marketing Communications*, *16*(3), 163-181.

Mohan, P., Padmanabhan, V. N., & Ramjee, R. (2008). Nericell: rich monitoring of road and traffic conditions using mobile smartphones. In: *Proceedings of the 6th ACM conference on Embedded network sensor systems* (pp. 323-336). ACM.

Moniz, A. (2015). Inferring employees' social media perceptions of goal-setting corporate cultures and the link to firm value. *Unpublished Working Paper*.

Moniz, A., & de Jong, F. (2014). Sentiment analysis and the impact of employee satisfaction on firm earnings. In: *European Conference on Information Retrieval* (pp. 519-527). Springer, Cham.

Monopolkommission (2018). Wettbewerb 2018, XXII. Hauptgutachten der Monopolkommission gemäß § 44 Abs. 1 Satz 1 GWB. Retrieved from: https://www.monopolkommission.de/de/gutachten/hauptgutachten/212-xxii-gesamt.html (accessed December 29, 2019).

Moore, T., Clayton, R., & Anderson, R. (2009). The economics of online crime. *Journal of Economic Perspectives*, *23*(3), 3-20.

Müller, J., Kreuz, S., Höhl, W., & Lüdecke, V. (2017, October). A Process Full of Challenges: A Serious Game About the German Energy Transition. In: *European Conference on Games Based Learning* (pp. 456-463). Academic Conferences International Limited.

Muslukhov, I., Boshmaf, Y., Kuo, C., Lester, J., & Beznosov, K. (2012). Understanding users' requirements for data protection in smartphones. In: *2012 IEEE 28th International Conference on Data Engineering Workshops* (pp. 228-235). IEEE.

Nikiforakis, N., & Slonim, R. (2019). Editors' Preface: Trends in experimental economics (1975–2018). *Journal of the Economic Science Association*, *5*, 1-6.

Nissenbaum, H. (2009). *Privacy in context: Technology, policy, and the integrity of social life*. Stanford University Press.

Niu, C., Zheng, Z., Wu, F., Tang, S., Gao, X., & Chen, G. (2018). Unlocking the value of privacy: Trading aggregate statistics over private correlated data. In: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2031-2040. ACM.

Norberg, P. A., Horne, D. R., & Horne, D. A. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of Consumer Affairs*, *41*(1), 100-126.

Nordhaus, W. (2019). Climate change: the ultimate challenge for economics. *American Economic Review*, *109*(6), 1991-2014.

Nunley, J. M., Owens, M. F., & Howard, R. S. (2011). The effects of information and competition on racial discrimination: Evidence from a field experiment. *Journal of Economic Behavior & Organization*, *80*(3), 670-679.

Olson M. (1965). *The logic of collective action—public goods and the theory of groups*. Harvard University Press, Cambridge.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251).

O'Sullivan, A., Henrick, B., Dixon, B., Barile, D., Zivkovic, A., Smilowitz, J., Lemay, D., Martin, W., German, J. B. & Schaefer, S. E. (2018). 21st century toolkit for optimizing population health through precision nutrition. *Critical Reviews in Food Science and Nutrition*, *58*(17), 3004-3015.

Oswald, A., & Stern, N. (2019). Why does the economics of climate change matter so much, and why has the engagement of economists been so weak?. *Royal Economic Society Newsletter*. Online: http://andrewoswald.com/docs/ClimatechangeOswaldSternSept2019forRES.pdf (accessed April 04, 2021)

Palser, E. R., Lazerwitz, M., & Fotopoulou, K. (2021). Gender and geographical disparity in editorial boards of journals in psychology and neuroscience. *bioRxiv*.

Park, E. S. (2000). Warm-glow versus cold-prickle: a further experimental study of framing effects on free-riding. *Journal of Economic Behavior & Organization*, *43*(4), 405-421.

Payback (2020). Facts and Figures. Online: https://www.payback.net/en/about-payback/facts-figures/ (accessed October 08, 2020).

Persson, A. J., & Hansson, S. O. (2003). Privacy at work–ethical criteria. *Journal of Business Ethics*, *42*(1), 59-70.

Pfeffer, J., Zorbach, T., & Carley, K. M. (2014). Understanding online firestorms: Negative word-of-mouth dynamics in social media networks. *Journal of Marketing Communications*, *20*(1-2), 117-128.

Plesch, J., & Wolff, I. (2018). Personal-data disclosure in a field experiment: Evidence on explicit prices, political attitudes, and privacy preferences. *Games*, *9*(2), 24.

Poch, R., & Martin, B. (2015). Effects of intrinsic and extrinsic motivation on user-generated content. *Journal of Strategic Marketing*, *23*(4), 305-317.

Poutanen, K., Nordlund, E., Paasi, J., Vehmas, K., & Åkerman, M. (2017). Food economy 4.0: VTT's vision towards intelligent, consumer-centric food production.

Quercia, D., Las Casas, D. B., Pesce, J. P., Stillwell, D., Kosinski, M., Almeida, V. A., & Crowcroft, J. (2012). Facebook and Privacy: The Balancing Act of Personality, Gender, and Relationship Currency. *Proceedings of the International AAAI Conference on Web and Social Media*, *6*(1), 306-313.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review*, *83*(5), 1281-1302.

Ramírez-Correa, P., Grandón, E. E., & Rondán-Cataluña, F. J. (2020). Users segmentation based on the Technological Readiness Adoption Index in emerging countries: The case of Chile. *Technological Forecasting and Social Change*, *155*, 120035.

Reuter, N., & Junge, C. (2017). Arbeitgeberbewertungsportale – die neue Macht der Bewerber?. In: *Praxishandbuch Social Media Recruiting* (pp. 277-323). Springer Gabler, Wiesbaden.

Rochedo, P. R., Soares-Filho, B., Schaeffer, R., Viola, E., Szklo, A., Lucena, A. F., Koberle, A., Davis, J. L., Rajao, R. & Rathmann, R. (2018). The threat of political bargaining to climate mitigation in Brazil. *Nature Climate Change*, *8*(8), 695-698.

Rompf, S., Kroneberg, C., & Schlösser, T. (2017). Institutional trust and the provision of public goods: When do individual costs matter? The case of recycling. *Rationality and Society*, *29*(2), 160-178.

Roth, A. E. (1993). On the early history of experimental economics. *Journal of the History of Economic Thought*, *15*(2), 184-209.

Rowe, F. (2020). Contact tracing apps and values dilemmas: A privacy paradox in a neo-liberal world. *International Journal of Information Management*, *55*, 102178.

Roy, S. K., Balaji, M. S., Sadeque, S., Nguyen, B., & Melewar, T. C. (2017). Constituents and consequences of smart customer experience in retailing. *Technological Forecasting and Social Change*, *124*, 257-270.

Sauerland, M., & Höhs, J. (2019). Reden ist Silber, Schweigen ist Geld? – Tabuthema Geld. In: *Geld-Vom Sein zum Schein* (pp. 37-63). Springer, Wiesbaden.

Schleyer-Lindenmann, A., Ittner, H., Dauvier, B., & Piolat, M. (2018). Die NEP-Skala–hinter den (deutschen) Kulissen des Umweltbewusstseins. *Diagnostica*, *64*(3), 156-167.

Schmuck, D., Matthes, J., & Naderer, B. (2018). Misleading consumers with green advertising? An affect–reason–involvement account of greenwashing effects in environmental advertising. *Journal of Advertising*, *47*(2), 127-145.

Schudy, S., & Utikal, V. (2018). Does imperfect data privacy stop people from collecting personal data?. *Games*, *9*(1), 14.

Schudy, S., & Utikal, V. (2017). 'You must not know about me'—On the willingness to share personal data. *Journal of Economic Behavior & Organization*, *141*, 1-13.

Schultz, P. W. (1999). Changing behavior with normative feedback interventions: A field experiment on curbside recycling. *Basic and Applied Social Psychology*, *21*(1), 25-36.

Schwarzmüller, T., Brosi, P., Duman, D., & Welpe, I. M. (2018). How does the digital transformation affect organizations? Key themes of change in work design and leadership. *mrev management revue*, *29*(2), 114-138.

Selten, R. (1967). Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments. In: Sauermann, H. (ed.), *Beiträge zur Experimentellen Wirtschaftsforschung* (pp. 136–168). JCB Mohr (Paul Siebeck): Tübingen, Germany.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Wadsworth Cengage Learning.

Shirin, A., & Kleyn, N. (2017). An evaluation of the effects of corporate reputation on employee engagement: The case of a major bank in South Africa. *International Studies of Management & Organization*, *47*(3), 276-292.

Smith, H. J., Dinev, T., & Xu, H. (2011). Information privacy research: an interdisciplinary review. *MIS Quarterly*, *35*(4), 989-1015.

Smith V. L. (2010). Experimental methods in economics. In: Durlauf S. N., & Blume L. E. (eds.) *Behavioural and Experimental Economics* (pp. 120-136). The New Palgrave Economics Collection. Palgrave Macmillan, London

Smith, V. L. (1991). *Papers in experimental economics*. Cambridge University Press.

Son, J. Y., & Kim, S. S. (2008). Internet users' information privacy-protective responses: A taxonomy and a nomological model. *MIS Quarterly*, *3*, 503-529.

Sonnemans, J., Schram, A., & Offerman, T. (1998). Public good provision and public bad prevention: The effect of framing. *Journal of Economic Behavior & Organization*, *34*(1), 143-161.

Spiekermann, S., Grossklags, J., & Berendt, B. (2001). E-privacy in 2nd generation E-commerce: privacy preferences versus actual behavior. In: *Proceedings of the 3rd ACM conference on Electronic Commerce* (pp. 38-47). ACM.

Splendid Research (2019). Optimized Self Monitor 2019. Online: https://www.splendid-research.com/de/studie-optimized-self.html (accessed October 20, 2020).

Stewart-Knox, B., Kuznesof, S., Robinson, J., Rankin, A., Orr, K., Duffy, M., Poínhos, R., de Almeida, M. D. V., Macready, A., Gallagher, C., Berezowska, A., Fischer, A. R. H., Navas-Carretero, S., Riemer, M., Traczyk, I., Gjelstad, I. M. F., Mavrogianni, C. & Frewer, L. J. (2013). Factors influencing European consumer uptake of personalised nutrition. Results of a qualitative analysis. *Appetite*, *66*, 67-74.

Svorenčík, A., & Maas, H. (2016). *The making of experimental economics: Witness seminar on the emergence of a field*. Springer.

Taddicken, M. (2014). The 'privacy paradox'in the social web: The impact of privacy concerns, individual characteristics, and the perceived social relevance on different forms of self-disclosure. *Journal of Computer-Mediated Communication*, *19*(2), 248-273.

Tadelis, S. (2016). Reputation and feedback systems in online platform markets. *Annual Review of Economics*, *8*, 321-340.

Tavoni, A., Dannenberg, A., Kallis, G., & Löschel, A. (2011). Inequality, communication, and the avoidance of disastrous climate change in a public goods game. *Proceedings of the National Academy of Sciences*, *108*(29), 11825-11829.

Techniker Krankenkasse (2020). So funktioniert TK-Fit. Online: https://www.tk.de/techniker/magazin/digitale-gesundheit/tkf/tk-fit/belohnungen-fitnessprogramm-2066246 (accessed October 12, 2020).

Tedjamulia, S. J., Dean, D. L., Olsen, D. R., & Albrecht, C. C. (2005). Motivating content contributions to online communities: Toward a more comprehensive theory. In: *Proceedings of the 38th annual Hawaii international conference on system sciences* (pp. 193b-193b). IEEE.

The Nobel Foundation (2019). The Prize in Economic Sciences 2019. NobelPrize.org. Online: https://www.nobelprize.org/prizes/economic-sciences/2019/summary/ (accessed April 04, 2021).

The Nobel Foundation (2018). The Prize in Economic Sciences 2018. NobelPrize.org. Online: https://www.nobelprize.org/prizes/economic-sciences/2018/summary/ (accessed April 04, 2021).

The Nobel Foundation (2017). The Prize in Economic Sciences 2017. NobelPrize.org. Online: https://www.nobelprize.org/prizes/economic-sciences/2017/summary/ (accessed April 04, 2021).

The Nobel Foundation (2002). The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2002. NobelPrize.org. Online: https://www.nobelprize.org/prizes/economic-sciences/2002/summary/ (accessed April 04, 2021).

Tietenberg, T. H., & Lewis, L. (2018). *Environmental and natural resource economics*. 11th Edition, Routledge.

Tresp, V., Overhage, J. M., Bundschus, M., Rabizadeh, S., Fasching, P. A., & Yu, S. (2016). Going digital: a survey on digitalization and large-scale data analytics in healthcare. *Proceedings of the IEEE*, *104*(11), 2180-2206.

Trudel, R. (2019). Sustainable consumer behavior. *Consumer Psychology Review*, *2*(1), 85-96.

Trudel, R., Argo, J. J., & Meng, M. D. (2016). Trash or recycle? How product distortion leads to categorization error during disposal. *Environment and Behavior*, *48*(7), 966-985.

Tsai, J. Y., Egelman, S., Cranor, L., & Acquisti, A. (2011). The effect of online privacy information on purchasing behavior: An experimental study. *Information Systems Research*, *22*(2), 254-268.

Turban, D. B., & Cable, D. M. (2003). Firm reputation and applicant pool characteristics. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, *24*(6), 733-751.

Turow, J., Feldman, L., & Meltzer, K. (2005). Open to exploitation: America's shoppers online and offline. *Departmental Papers (ASC)*, 35.

United Nations (2015). Paris Agreement.

Uzhegova, M., Torkkeli, L., & Saarenketo, S. (2019). Corporate Social Responsibility in SMEs: Implications on Competitive. *mrev management revue*, *30*(2-3), 232-267.

Vasilaky, K. N., & Brock, J. M. (2020). Power(ful) guidelines for experimental economists. *Journal of the Economic Science Association*, *6*(2), 189-212.

Vitak, J., Liao, Y., Kumar, P., Zimmer, M., & Kritikos, K. (2018). Privacy attitudes and data valuation among fitness tracker users. In: *International Conference on Information* (pp. 229-239). Springer, Cham.

Wahlström, J., Skog, I., & Händel, P. (2017). Smartphone-based vehicle telematics: A ten-year anniversary. *IEEE Transactions on Intelligent Transportation Systems*, *18*(10), 2802-2825.

Waichman, I., Requate, T., Karde, M., & Milinski, M. (2018). *Challenging conventional wisdom: Experimental evidence on heterogeneity and coordination in avoiding a collective catastrophic event* (No. 2018-05). Economics Working Paper.

Wang, R. J. H., Krishnamurthi, L., & Malthouse, E. C. (2018). When reward convenience meets a mobile app: Increasing customer participation in a coalition loyalty program. *Journal of the Association for Consumer Research*, *3*(3), 314-329.

Wayne, J. H., & Casper, W. J. (2012). Why does firm reputation in human resource policies influence college students? The mechanisms underlying job pursuit intentions. *Human Resource Management*, *51*(1), 121-142.

Weimann, J., & Brosig-Koch, J. (2019). *Einführung in die experimentelle Wirtschaftsforschung*. Springer Berlin Heidelberg.

Williamson, B. (2015). Algorithmic skin: Health-tracking technologies, personal analytics and the biopedagogies of digitized health and physical education. *Sport, Education and Society*, *20*(1), 133-151.

Wright, D., Finn, R., Gellert, R., Gutwirth, S., Schütz, P., Friedewald, M., Venier, S. & Mordini, E. (2014). Ethical dilemma scenarios and emerging technologies. *Technological Forecasting and Social Change*, *87*, 325-336.

Ye, S., Gao, G., & Viswanathan, S. (2014). Strategic Behavior in Online Reputation Systems. *MIS Quarterly*, *38*(4), 1033-1056.

Zelmer, J. (2003). Linear public goods experiments: A meta-analysis. *Experimental Economics*, *6*(3), 299-310.

Zhang, L., & Ortmann, A. (2013). Exploring the meaning of significance in experimental economics. *UNSW Australian School of Business Research Paper* 2013-32.

Zheng, J., Gao, D. W., & Lin, L. (2013). Smart meters in smart grid: An overview. In: *2013 IEEE Green Technologies Conference (GreenTech)* (pp. 57-64). IEEE.